

EDITED BY

MARK D.
WHITE



≡ The Oxford Handbook of
**ETHICS AND
ECONOMICS**

THE OXFORD HANDBOOK OF

**ETHICS AND
ECONOMICS**

THE OXFORD HANDBOOK OF

ETHICS AND
ECONOMICS

Edited by

MARK D. WHITE

OXFORD
UNIVERSITY PRESS

OXFORD

UNIVERSITY PRESS

Great Clarendon Street, Oxford, OX2 6DP,
United Kingdom

Oxford University Press is a department of the University of Oxford.
It furthers the University's objective of excellence in research, scholarship,
and education by publishing worldwide. Oxford is a registered trade mark of
Oxford University Press in the UK and in certain other countries

© Oxford University Press 2019

The moral rights of the authors have been asserted

First Edition published in 2019

Impression: 1

All rights reserved. No part of this publication may be reproduced, stored in
a retrieval system, or transmitted, in any form or by any means, without the
prior permission in writing of Oxford University Press, or as expressly permitted
by law, by licence or under terms agreed with the appropriate reprographics
rights organization. Enquiries concerning reproduction outside the scope of the
above should be sent to the Rights Department, Oxford University Press, at the
address above

You must not circulate this work in any other form
and you must impose this same condition on any acquirer

Published in the United States of America by Oxford University Press
198 Madison Avenue, New York, NY 10016, United States of America

British Library Cataloguing in Publication Data
Data available

Library of Congress Control Number: 2019941475

ISBN 978-0-19-879399-1

Printed and bound by
CPI Group (UK) Ltd, Croydon, CRO 4YY

Links to third party websites are provided by Oxford in good faith and
for information only. Oxford disclaims any responsibility for the materials
contained in any third party website referenced in this work.

ACKNOWLEDGMENTS

I would like to thank Adam Swallow at Oxford University Press for his indefatigable support and help with getting this project off the ground and completed; his associate Katie “Hawkeye” Bishop for her invaluable editorial assistance; and all of the contributors to this volume for their fantastic chapters and cooperation throughout the process.

For their unending inspiration, I extend special gratitude to Amartya Sen, Amitai Etzioni, and Deirdre McCloskey.

TABLE OF CONTENTS

<i>List of Figures</i>	xi
<i>List of Tables</i>	xiii
<i>Notes on Contributors</i>	xv

Introduction	1
MARK D. WHITE	

PART I FOUNDATIONS

A. ETHICAL THEORIES IN ECONOMICS

1. Adam Smith and the Study of Ethics in a Commercial Society	13
STEFANIE HAEFFELE AND VIRGIL HENRY STORR	
2. Virtue and Economics, Horse and Cart	34
JENNIFER A. BAKER	
3. With All Due Respect: A Kantian Approach to Economics	54
MARK D. WHITE	
4. Ethical Pluralism in Economics	77
JONATHAN B. WIGHT	
5. Economic Ethics and the Capability Approach	96
CONSTANZE BINDER AND INGRID ROBEYNS	

B. SOURCES OF MORALITY IN ECONOMICS

6. Evolution and Moral Motivation in Economics	117
GEOFFREY M. HODGSON	
7. Morality as a Complex Adaptive System: Rethinking Hayek's Social Ethics	138
GERALD GAUS	

8. On the Evolution of Ethics, Rationality, and Economic Behavior 160
DAVID C. ROSE

C. ETHICS IN SCHOOLS OF ECONOMICS

9. Human Ethicality: Evidence and Insights from Behavioral Economics 181
SANJIT DHAMI AND ALI AL-NOWAIHI
10. Ethics *and* Economics: A Complex Systems Approach 208
JOHN B. DAVIS
11. Economics and Ethics within the Austrian School of Economics 229
PETER J. BOETTKE AND KAITLYN WOLTZ
12. Feminist Economics and Ethics 248
ULRIKE KNOBLOCH
13. Economy and Culture: The Importance of Sense-Making 270
ARJO KLAMER

PART II APPLICATIONS

A. COMMERCE AND MARKETS

14. Humane Markets: The Classical Tradition of Political Economy 295
JAMES R. OTTESON
15. Capitalism and Democracy: Allies, Rivals, or Strangers? 316
JULIAN REISS
16. The Moral Status of Profit 337
JOSEPH HEATH
17. The Ethics of Money and Finance 358
JOAKIM SANDBERG
18. Ethics *And, In, and For* Labor Markets 381
MICHAEL S. MCPHERSON AND DEBRA SATZ

B. WELFARE, RISK, AND POLICY

19. Cost-Benefit Analysis and Social Welfare Functions 399
MATTHEW D. ADLER

20. The Normative Economics of Social Risk	423
MARC FLEURBAEY	
21. The Ethics of Making Risky Decisions for Others	446
LUC BOVENS	
22. The Tragedy of Economics: On the Nature of Economic Harm and the Responsibilities of Economists	474
GEORGE F. DEMARTINO	

C. ETHICS IN APPLIED ECONOMICS

23. Economics, Ethics, and Health Insurance	499
DANIEL M. HAUSMAN	
24. Deontological Morality and Economic Analysis of Law	521
EYAL ZAMIR AND BARAK MEDINA	
25. The Ethics and Economics of Ecological Justice	544
DAVID SCHMIDTZ	
26. Civil Rights, Employment, and Race	559
BRENDAN O'FLAHERTY	

PART III CONCLUSION

27. Lessons from Economics	583
JOHN BROOME	
<i>Index</i>	607

LIST OF FIGURES

4.1	Multiple approaches to understanding right from wrong	79
9.1	Results of the baseline treatment in the die-cup experiment	185
12.1	Shifting processes of work in the four sector-model	261
13.1	The model of five spheres	272
16.1	Supply and demand	344
16.2	Exercising market power	345
16.3	Competition among suppliers	347
16.4	Direct normative prescription vs. adversarialism	353
27.1	Contours of equal goodness	601

LIST OF TABLES

9.1	Payoffs of the players in each treatment in Gneezy (2005)	184
10.1	The mainstream and social economic visions of ethics <i>and</i> economics	213
10.2	The mainstream economic vision: Society embedded in the economy	218
10.3	Four different types of individual advantage and associated moral values	222
10.4	Positions regarding the pervasiveness of moral values in economic life	224
19.1	Types of well-being comparisons	406
19.2	Invariance to rescalings of the well-being measure	408
20.1	Utilitarianism vs. fairness	424
20.2	Do ex post inequalities matter?	424
20.3	Enriched description of consequences	428
20.4	Pareto vs. inequality aversion	429
20.5	Ex post egalitarianism vs. separability	434
20.6	Implications of separability	435
20.7	Pareto for equal risk, separability and statewise dominance vs. inequality aversion	435
20.8	Prevention vs. vaccination	436
21.1	Risky	451
21.2	Certain	451
21.3	Distributional Sensitivities, Prospect Values, and Parameters	454
21.4	<i>Restrictive Intervention</i> and <i>Intervention for All</i>	455
21.5	<i>Single-Cause</i> and <i>Multiple-Cause</i>	456
21.6	<i>No Screening</i> and <i>Screening</i>	457
21.7	<i>Vaginal Birth</i> and <i>C-section</i>	458
21.8	Overview	460
21.9	<i>Prevention</i> and <i>Treatment</i>	461
21.10	<i>Protecting 99 Sheep</i> and <i>Saving One Sheep</i>	462
21.11	<i>Safe Detonation</i> and <i>Lethal Bomb</i>	463

21.12	<i>One Lives One Dies</i> and <i>Coin Flip</i>	465
21.13	Adler's Objection	466
21.14	Utilities in Otsuka and Voorhoeve's Example	467
21.15	Otsuka and Voorhoeve's Objection to Prioritarianism	468
21.16	Sen's Libertarian Paradox	469
21.17	Gibbard's version of the Libertarian Paradox	469
22.1	Compensable vs. noncompensable harm	482
22.2	A taxonomy of harmed or harmful conditions	484
26.1	Results of an employment audit	570

NOTES ON CONTRIBUTORS

Matthew D. Adler is the Richard A. Horvitz Professor of Law and Professor of Economics, Philosophy and Public Policy at Duke University. He was previously the Leon Meltzer Professor of Law at the University of Pennsylvania School of Law. His scholarship is interdisciplinary, drawing from both welfare economics and normative ethics. Adler's current research focuses on the theoretical foundations and practical implementation of prioritarianism. With Ole Norheim, he is the founder of the "Prioritarianism in Practice" research network. Adler is the author of *New Foundations of Cost-Benefit Analysis* (Harvard, 2006, with Eric Posner); *Well-Being and Fair Distribution: Beyond Cost-Benefit Analysis* (Oxford, 2012), and *Measuring Social Welfare* (Oxford, 2019), and is the editor (with Marc Fleurbaey) of the *Oxford Handbook of Well-Being and Public Policy* (2016). Adler was until 2017 an editor of the journal *Legal Theory*, and is currently an editor of *Economics and Philosophy*.

Ali al-Nowaihi is Emeritus Professor of Economics in the School of Business at Leicester University. Since 2000, his main area of research has been in behavioral economics. His main coauthor in this enterprise has been Sanjit Dhami; he has also published with Livio Stracca, Andrew Colman, Briony Pulford, David Omtzigt, Maxine Wei, and Cass Sunstein. Other areas he has published in include algebra, oligopoly theory (with Paul Levine), monetary policy (with Paul Levine and with Sanjit Dhami), and public economics (with Clive Fraser and with Sanjit Dhami). Ali graduated in mathematics from Cairo University in 1971, obtained his MSc also in mathematics from London University in 1973, and obtained his PhD in oligopoly theory in 1976 from the South Bank Polytechnic (now University). Ali has taught in two secondary schools, the South Bank Polytechnic (now University), Thames Polytechnic (now Greenwich University), and at Leicester University. He has also supervised six successful PhD students.

Jennifer A. Baker is Professor of Philosophy at the College of Charleston. Her focus is on updating ancient virtue ethics for use today. Her published articles include "Who is Afraid of a Final End? The Omission of Traditional Practical Rationality from Contemporary Virtue Ethics," "Virtue Ethics and Practical Guidance," and "Virtue and Behavior." Her co-edited collection of work on virtue and economics (with Mark D. White) was published by Oxford University Press in 2016. She is at work applying Stoic virtue ethics to business ethics as well as current-day policing.

Constanze Binder is Assistant Professor in Philosophy at the Faculty of Philosophy, Co-Director of the Erasmus Institute for Philosophy and Economics, and Programme Director of the Research Master programme in Philosophy and Economics at Erasmus

University Rotterdam. Constanze studied economics and environmental system science at Graz University and obtained a PhD in philosophy at the University of Groningen. She previously taught in the Philosophy Departments of Groningen and Leiden University, worked at the Economics Department of the University of Osnabrück, and contributed to projects for the Austrian Federal Ministry of the Environment and the Austrian Human Dimensions Program. Constanze's research is on the interface of philosophy and economics, with a particular focus on the analysis of freedom, responsibility, and distributive justice in political philosophy and welfare economics, as well as on the ethics of individual and collective decision-making in politics and economics.

Peter J. Boettke is University Professor of Economics and Philosophy at George Mason University and Director of the F.A. Hayek Program for Advanced Study in Philosophy, Politics, and Economics at the Mercatus Center. He received his PhD in economics from George Mason University in 1989. Before joining the faculty at George Mason University in 1998, Boettke taught at New York University. He currently serves as the President of the Mont Pelerin Society and was the President of the Southern Economics Association from 2015–17. His publications include *Why Perestroika Failed: The Politics and Economics of Socialist Transformation*, *Calculation and Coordination: Essays on Socialism and Transitional Political Economy*, *Challenging Institutional Analysis and Development: The Bloomington School*, *Robust Political Economy for the 21st Century*, and *Living Economics: Yesterday, Today and Tomorrow*.

Luc Bovens (PhD University of Minnesota) is a Professor of Philosophy at the University of North Carolina at Chapel Hill and is core faculty in its PPE program. He was a Professor of Philosophy in the University of Colorado at Boulder (1990–2003) and in the London School of Economics and Political Science (2004–2017). He is joint author of *Bayesian Epistemology* (Oxford, 2003). His areas of research are philosophy of economics, philosophy and public policy, formal epistemology, rationality, and moral psychology.

John Broome is Emeritus White's Professor of Moral Philosophy at the University of Oxford, and a fellow of Corpus Christi College, Oxford. He is also an Honorary Professor at the Australian National University. Before moving to Oxford, he was Professor of Philosophy at the University of St Andrews, and before that Professor of Economics at the University of Bristol. He works on rationality and normativity, and also on value theory. As an application of his work on value theory, he is involved in the philosophy of climate change. He was a Lead Author of the *Fifth Assessment Report of the Intergovernmental Panel on Climate Change*. His books include, on value theory, *Weighing Goods*, *Weighing Lives*, and *Ethics Out of Economics*; on climate change, *Climate Matters*; and on rationality and normativity, *Rationality Through Reasoning*.

John B. Davis is Professor Emeritus of Economics, Marquette University, USA, and Professor Emeritus of Economics, University of Amsterdam, Netherlands. He is the author of *Keynes's Philosophical Development* (Cambridge, 1994), *The Theory of the Individual in Economics* (Routledge, 2003), *Individuals and Identity in Economics* (Cambridge, 2011), co-author with Marcel Boumans of *Economic Methodology: Understanding Economics*

as a Science (Palgrave Macmillan, 2010), and co-author with Robert McMaster of *Health Care Economics* (Routledge, 2017). He is a former editor of the *Review of Social Economy*, is co-editor with Wade Hands of the *Journal of Economic Methodology*, and is the editor of the Routledge Advances in Social Economics book series.

George F. DeMartino is Professor of Economics in the Josef Korbel School of International Studies, University of Denver, where he co-directs the MA in Global Finance, Trade and Economic Integration. He has published *Global Economy, Global Justice: Theoretical Objections and Policy Alternatives to Neoliberalism* (Routledge) and *The Economist's Oath: On the Need for and Content of Professional Economic Ethics* (Oxford University Press). He is co-editor with Deirdre N. McCloskey of *The Oxford Handbook of Professional Economic Ethics*. He is now working on *The Tragedy of Economics: Harm, Economic Harm, and the Harm Economists Do as They Try to Do Good*.

Sanjit Dhami is Professor of Economics at the University of Leicester. He studied at the Delhi School of Economics for his MPhil and the University of Toronto for his Masters and PhD degrees in economics. He has previously taught at the Universities of Toronto, Essex, and Newcastle. His research has mainly focused on behavioral economic theory and its applications. He is the author of *Foundations of Behavioral Economic Analysis* published by Oxford University Press in 2016, possibly the most comprehensive graduate treatment of the subject. His research spans several areas in behavioral economics, which include but are not restricted to behavioral decision theory, other-regarding preferences, time preferences, behavioral game theory, and judgment heuristics.

Marc Fleurbaey is an economist, a professor in the Woodrow Wilson School and Center for Human Values at Princeton University, and a member of the Collège d'Études Mondiales (Paris FMSH). He is the co-author of *Beyond GDP* (with Didier Blanchet, Cambridge, 2013), *A Theory of Fairness and Social Welfare* (with François Maniquet, Cambridge, 2011), and the author of *Fairness, Responsibility and Welfare* (Oxford, 2008). He was a coordinating lead author for the IPCC 5th Report, and one of the initiators of the International Panel on Social Progress.

Gerald Gaus is the James E. Rogers Professor of Philosophy at the University of Arizona, where he directs the program in Philosophy, Politics, Economics, and Law. His books include *Value and Justification* (Cambridge, 1990), *Justificatory Liberalism* (Oxford, 1996), and *The Order of Public Reason* (Cambridge, 2011). His most recent book is *The Tyranny of the Ideal: Justice in a Diverse Society* (Princeton University Press, 2016). He is currently writing a book on morality, complexity, and evolution to be published by Oxford University Press. His papers can be found at www.gaus.biz.

Stefanie Haeffele is a Senior Research Fellow, the Deputy Director of Academic and Student Programs, and a Senior Fellow with the F.A. Hayek Program for Advanced Study in Philosophy, Politics, and Economics at the Mercatus Center at George Mason University. She is the co-author of *Community Revival in the Wake of Disaster: Lessons in Local Entrepreneurship* (Palgrave), and the editor of *Knowledge and Incentives in*

Policy: Using Public Choice and Market Process Theory to Analyze Public Policy Issues (Rowman and Littlefield International).

Daniel M. Hausman earned his philosophy PhD in 1978 at Columbia University, and has taught at the University of Maryland at College Park, Carnegie Mellon University, and since 1988 at the University of Wisconsin-Madison; he has visited at the Institute for Advanced Studies, the London School of Economics, and the Université de Cergy-Pontoise. His research focuses on issues at the boundaries between economics and philosophy. With Michael McPherson, he founded the journal *Economics and Philosophy*. He is the editor of *The Philosophy of Economics: An Anthology* (3rd edn, 2007). His books include *Capital, Profits, and Prices*, *The Inexact and Separate Science of Economics*, *Causal Asymmetries*, *Preference, Value, Choice and Welfare*, *Valuing Health: Well-being, Freedom, and Suffering*, and *Economic Analysis, Moral Philosophy, and Public Policy* (3rd edn, 2017, co-authored with Michael McPherson and Debra Satz). In 2009, Daniel Hausman was elected to the American Academy of Arts and Sciences.

Joseph Heath is Professor in the Department of Philosophy and the School of Public Policy and Governance at the University of Toronto. A fellow of the Royal Society of Canada and the Trudeau Foundation, Heath is the author of several books, both popular and academic. His most recent, *Morality, Competition and the Firm* (Oxford, 2014), is a collection of papers on business ethics and the normative foundations of market economies. He is also the author of *Enlightenment 2.0*, which won the Shaughnessy Cohen Prize for Political Writing in 2015.

Geoffrey M. Hodgson is Professor in Management at Loughborough University London. He is the author of several books, including *Conceptualizing Capitalism* (2015, winner of the Schumpeter Prize), *From Pleasure Machines to Moral Communities* (2013), *Darwin's Conjecture* (with Thorbjørn Knudsen, 2010), and *How Economics Forgot History* (2001). He is the author of 150 articles in academic journals, Editor-in-Chief of the *Journal of Institutional Economics*, and Secretary of the World Interdisciplinary Network for Institutional Research.

Arjo Klammer is a professor of cultural economics at Erasmus University in the Netherlands. He is the past president of the Association of Cultural Economists. He has developed a value-based approach and introduced the notion of shared goods. Among his publications are *Doing the Right Thing: A Value Based Economy* (2017) and *Speaking of Economics: How to Be in the Conversation* (2007).

Ulrike Knobloch is Assistant Professor of Economics and Gender at the University of Vechta, Germany, and lecturer at the University of Fribourg, Switzerland, and the University of Gießen, Germany. She holds a PhD in economics from the University of St. Gallen, Switzerland, as well as a diploma in economics and an intermediate diploma in philosophy from the University of Freiburg, Germany. She is a long-time active member of the following networks: *efas* – economics, feminism and science; European Platform of Women Scientists (EPWS); International Association for Feminist Economics

(IAFFE); and Network Caring Economy (which she co-founded in 1992). Her academic research focuses on the normative foundations of pluralist feminist economics, feminist economic ethics, regulatory ethics from a gender perspective, economics of provisioning and systems of provisioning in comparison, feminist critical thinking in social economy, and household economics. Her most recent publications include the edited volume *Economics of Provisioning: Contributions to the Plurality of Feminist Economic Theory* (2019 in German; English translation planned).

Michael S. McPherson served as President of the Spencer Foundation for fourteen years before retiring in 2017. Earlier he was President of Macalester College in St. Paul, Minnesota, for seven years. He is a nationally known economist whose expertise focuses on the interplay between education and economics. McPherson is co-author or editor of several books, including *Lesson Plan: An Agenda for Change in American Higher Education*, *Crossing the Finish Line: Completing College at America's Public Universities*, *The Student Aid Game*, and *Economic Analysis, Moral Philosophy, and Public Policy* (with Daniel Hausman and Debra Satz). McPherson was founding co-editor of the journal *Economics and Philosophy* (with Hausman). He is currently a Senior Research Fellow at the Mellon Foundation and a non-resident Fellow at the Urban Institute.

Barak Medina is the Justice Haim H. Cohn Professor of Human Rights at the Faculty of Law, The Hebrew University of Jerusalem. He currently serves as Rector of the Hebrew University, after serving as Dean of the Faculty of Law from 2009 to 2012. His scholarly work focuses on human rights law and economic analysis of law. He has served as a visiting professor at the law schools of Columbia University and University of California, Berkeley.

Brendan O'Flaherty is Professor of Economics at Columbia University. Most of his work is about homelessness, crime, race, and Newark, New Jersey. His books include *Shadows of Doubt: Stereotypes, Crime, and the Pursuit of Justice* (with Rajiv Sethi), *The Economics of Race in the United States*, *How to House the Homeless* (with Ingrid Gould Ellen), *City Economics*, and *Making Room: The Economics of Homelessness*.

James R. Otteson is Thomas W. Smith Presidential Chair in Business Ethics and Professor of Economics at Wake Forest University. He received his BA from the University of Notre Dame and his PhD from the University of Chicago. He specializes in political economy, moral philosophy, and the history of economic thought, and his publications include *Adam Smith's Marketplace of Life* (2002), *Actual Ethics* (2006), *Adam Smith* (2013), and *The End of Socialism* (2014). His most recent book, *Honorable Business: A New Framework for Business in a Just and Humane Society*, was published by Oxford University Press in 2019.

Julian Reiss is Professor of Philosophy at Durham University and co-director of the Centre for Humanities Engaging Science and Society (CHESS). He has a degree in economics and finance from the University of St. Gallen and a PhD in philosophy from the London School of Economics. His main research interests are methodologies of the

sciences (especially causality and causal inference, models, simulations and thought experiments, and counterfactuals), philosophy of economics, and science and values. He is the author of *Error in Economics: Towards a More Evidence-Based Methodology* (2008), *Philosophy of Economics: A Contemporary Introduction* (2013), *Causation, Evidence, and Inference* (2015), and over fifty papers in leading philosophy and social science journals and edited collections.

Ingrid Robeyns is an economist and philosopher and holds the Chair in Ethics of Institutions at the Ethics Institute at Utrecht University. She currently also serves as the president of the Human Development and Capability Association. She works primarily in normative political philosophy and applied ethics, in particular theories of justice, the evaluation of institutions, and the capability approach. In her recent book *Well-Being, Freedom and Social Justice: The Capability Approach Re-examined* (Open Book Publishers), she analyses how all the different strands in the capability approach relate to each other and can be thought to fit together under one generalized overarching structure. In applied ethics and applied philosophy, she has worked on questions of gender, parenthood, disability, basic income, welfare state arrangements, and ecological sustainability, among many more. More information and links to her publications can be found at www.ingridrobeyns.info.

David C. Rose is Professor of Economics at the University of Missouri-St. Louis. His research focuses on behavioral economics, political economy, the theory of the firm, and ethics. He has published scholarly articles on a wide range of topics. His book *The Moral Foundation of Economic Behavior* (Oxford, 2011) explored the role that moral beliefs play in the development and operation of free market societies. His more recent book *Why Culture Matters Most* (Oxford, 2019) explores how culture uniquely solves the most daunting obstacle to individual and collective human flourishing: individual rationality undermining the common good. His work has been supported by the Weldon Spring Foundation, the National Institute of Mental Health, the HFL Foundation, the Earhart Foundation, and the Templeton Foundation. He earned a bachelor's degree in economics from Missouri State University and a PhD in economics from the University of Virginia.

Joakim Sandberg is Director of the Financial Ethics Research Group at University of Gothenburg in Sweden. He is Professor of Economics and Finance from a Humanist Perspective at University of Groningen in the Netherlands, and Associate Professor of Practical Philosophy at University of Gothenburg. He is also affiliated with the Mistra Center for Sustainable Markets (Misum) at the Stockholm School of Economics. Joakim has a background in both economics (MBA in 2003) and philosophy (PhD in 2008). He was selected as Wallenberg Academy Fellow by the Royal Swedish Academy of Sciences in 2014.

Debra Satz is the Vernon R. and Lysbeth Warren Anderson Dean of the School of Humanities and Sciences, the Marta Sutton Weeks Professor of Ethics in Society, and Professor of Philosophy and (by courtesy) Political Science at Stanford University.

Her publications include *Why Some Things Should Not be for Sale: The Moral Limits of Markets* (2010) and *Economic Analysis, Moral Philosophy and Public Policy* (2017, with Michael McPherson and Daniel Hausman). She has coedited several books and is the author of numerous articles and the editor of the journal *Philosophy and Public Affairs*. She was elected a Fellow of the American Academy of Arts and Sciences in 2018.

David Schmidt is Kendrick Professor of Philosophy and Eller Chair of Service-Dominant Logic in the McGuire Center for Entrepreneurship at the University of Arizona. He is editor in chief of the journal *Social Philosophy & Policy*, author of *Elements of Justice*, and co-editor (with Dan Shahar) of *Environmental Ethics: What Really Matters, What Really Works*, published by Oxford University Press.

Virgil Henry Storr is a Senior Research Fellow, Vice President of Academic and Student Programs, and the Don C. Lavoie Senior Fellow in the F.A. Hayek Program in Philosophy, Politics, and Economics at the Mercatus Center at George Mason University, and an Associate Professor of Economics in the Department of Economics at George Mason University. He is the author of *Understanding the Culture of Markets* (Routledge).

Mark D. White is Professor and Chair of the Department of Philosophy at the College of Staten Island/CUNY, and a member of the doctoral faculty in economics at the Graduate Center of CUNY. He is the author of seven books, including *Kantian Ethics and Economics: Autonomy, Dignity, and Character*; editor or co-editor of many volumes, including *Economics and the Virtues: Building a New Moral Foundation* (with Jennifer A. Baker); and author of over sixty journal articles and book chapters in economics, philosophy, and law. He is series editor of *On Ethics and Economics* (Rowman and Littlefield International) and *Perspectives from Social Economics* (Palgrave Macmillan), and a co-founder of the blog *Economics and Ethics*.

Jonathan B. Wight is Professor of Economics and International Studies in the Robins School of Business at the University of Richmond. He is past president of the Association for Social Economics and author of *Ethics in Economics: An Introduction to Moral Frameworks* and the academic novel *Saving Adam Smith: A Tale of Wealth, Transformation, and Virtue*. He is co-author of *Teaching the Ethical Foundations and Economics*, and co-founded and edits the blog *Economics and Ethics*.

Kaitlyn Woltz is a third year PhD student at George Mason University in the Economics department. She is a PhD fellow with the Mercatus Center and a Graduate Fellow with the F.A. Hayek Program for Advanced Study in Philosophy, Politics, and Economics. Kaitlyn's research interests are in the fields of Austrian economics, institutional analysis, and the political economy of criminal justice. Her current project looks at the role that prison periodicals played in prisons and the criminal justice system during the mid-twentieth century.

Eyal Zamir is the Augusto Levi Professor of Commercial Law at the Faculty of Law, The Hebrew University of Jerusalem, where he served as Dean of the Faculty of Law from 2002 to 2005. His spheres of interest include economic and behavioral analyses of law,

law and normative ethics, and contract law and theory. He has been a visiting researcher or visiting professor at the law schools of Harvard, Yale, NYU, Georgetown, UCLA, and Zurich. Eyal has authored or edited sixteen books and published more than sixty articles and book chapters, including ones in *Columbia Law Review*, *Journal of Legal Studies*, *University of Chicago Law Review*, *California Law Review*, and *Virginia Law Review*.

INTRODUCTION

MARK D. WHITE

THE birth of modern economics is credited to a moral philosopher who wrote of both sympathy and self-interest, but nearly 250 years later the connection between ethics and economics has worn thin. In the last century especially, economics has adopted a scientific, positivistic, and empirical sheen, and its ethical foundations have been more difficult to discern, although they are still there. As long as it studies the choices, interactions, and welfare of human beings, economics cannot deny its roots in moral philosophy. At the most basic level, choice and well-being matter for a reason—and more to the point, human beings matter for a reason. Economists may take these things for granted when they go about their work, much as chemists take the existence of matter for granted. Unlike chemists, though, economists' attitudes toward the persons they study help determine the methods, results, and interpretations of their research, as well as the recommendations they make to business and governments based on it.

This handbook, and the chapters that comprise it, do not attempt to restore ethics to economics, for it was never truly gone. When buyers and sellers meet in markets to trade money for good and services, they are seeking their own self-interest—but they are also often behaving honestly and honorably, although standard economic models make it difficult to explain why. Policies, regulations, and laws are typically designed to maximize aggregate welfare, and this is certainly *an* ethical goal—but not the only one, and economics needs to make room for other possibilities. The contributors to this volume highlight the ethical concepts inherent in economic analysis, extend the ethical foundations of economics into new areas of moral philosophy, and apply these techniques to a wide range of economic topics.

If “ethics and economics” or “economics and ethics” can be considered a field unto itself, it is a bifurcated one at best. The most visible part is methodological and critical, with economists, philosophers, and other scholars challenging mainstream economics for its narrow focus on the maximization of utility or welfare to the neglect of concepts such as virtue, rights, dignity, or justice. For example, Amartya Sen introduced the concept of commitment into economics to challenge the singular pursuit of preference-satisfaction (1977) and questioned the priority of values with respect to the Pareto

criterion (1970). Amitai Etzioni (1987, 1988) asked if simple preference-satisfaction was sufficient to explain altruistic behavior and introduced concepts inspired by Kantian ethics to give the standard economic model more depth. Deirdre McCloskey emphasized the importance of virtue, not only in the history of economic thought (2008) but also the history of commerce and trade (2006, 2010, 2016). Given the critical stance of this kind of work, it is natural that heterodox economists have taken the lead in questioning the ethical presuppositions of mainstream economics, chief among them being social economists (Davis and Dolfsma 2015), feminist economists (Ferber and Nelson 2003), and Austrian economists (Boettke and Coyne 2015). Many of the participants in this volume have made substantial contributions of their own in this area, and their chapters herein advance their integral work.

The other part of the field is not as revolutionary and has therefore been less recognized, but is no less valuable. It consists of the wealth of research into ethical behavior on the part of mainstream economists using more modest revisions or reinterpretations of economic theory. A prime example is the voluminous research, both theoretical and empirical, on the economics of altruism, in which economists study and explain the incidence of charitable behavior by either expanding models to include other-regarding preferences or utilities without violating any of the axioms of choice theory (Bergstrom 1999; Fehr and Schmidt 1999) or positing different types or sources of utility to motivate self-regarding and altruistic behavior, a modest modification to the standard model (Margolis 1984; Etzioni 1988). Copious experimental evidence confirms that people often do not behave as the simplistic models of self-interest predict (Fehr and Schmidt 2006), and behavioral economists, who recommend more significant changes to the constrained preference-satisfaction model (while leaving its core intact), also work to explain this behavior with more elaborate and psychologically realistic models of choice (Dhami 2016: Part II). Evolutionary economists work to explain how the psychological foundations of economic behavior allowed early human beings to adapt to their environment, which reinforces the foundations of both mainstream and behavioral conceptions of choice (Bowles and Gintis 2011). Again, many of the chapters in this volume are written by pioneers in this area of economics and ethics, providing background, extensions, and applications of their seminal work.

These two parts of economics and ethics overlap, of course; commentary on technique need not be separate from application of it. Even the relatively mainstream work on altruism requires some changes to the standard economic model of choice, however modest. The more significant methodological critiques of economic technique and theory can—and should—be applied to topics of economic interest such as consumption and production theory and policymaking. Gary Becker (1976) famously expanded the domain of economics to topics previously thought to be sociological or legal—such as marriage, discrimination, and crime—but he did so strictly within the neoclassical framework, which neglects their inherent ethical dimensions. An expansion of the ethical foundation of economic theory is necessary to explain all the relevant facets of these phenomena and to predict behavior accurately. This goes for more “normal” economic activity as well, such as consumption and production, which have important

ethical aspects that are taken for granted at best (and ignored at worst). Also, heterodox and political economists who urge more attention given to inequality and injustice are often reformers at heart—like Jeremy Bentham, the implicit inspiration behind mainstream political economy—urging changes to economics to produce positive change in the real world.

Ethics is essential not only to economic theory, policy, and practice, but perhaps even more so to education. Very few economics students are exposed to ethics in any explicit manner, and the implicit ethical foundations of the economics they are taught are rarely brought into the light. Excellent resources exist, including books useful as texts for beginning students (Wight 2015) and advanced ones (Hausman, McPherson, and Satz 2016; Dutt and Wilber 2010), as well as handbooks for practicing scholars (Peil and van Staveren 2009, and the present volume), but the importance of ethics to economics must be recognized by the institutions responsible for education. To this end, movements on part of students and scholars, such as *Rethinking Economics* (Fischer et al. 2017), are prompting reconsideration of economics education around the globe (particularly since the Great Recession), and ethics needs to part of this in order for economics to reach its full potential. (The ethical behavior of economists themselves has also attracted increased attention in recent years; see DeMartino 2011 and DeMartino and McCloskey 2016.)

Neglect of ethics means ruling out possible explanations of behavior, unexamined topics of study, and crucial perspectives on policy issues. Economics cannot claim to speak to all aspects of life while maintaining an insufficiently broad perspective that obscures important aspects of the world. It can be a worthy heuristic experiment to see how much an assumption can explain—and the mainstream paradigm certainly explains a lot—but when it is clear the assumption fails, there is no virtue in defending it as if it had intrinsic rather than instrumental value. The methods and approaches of economics are tools to explain and predict human behavior, design institutions, and recommend policy, and economists should be willing to modify, extend, or reject them when sufficient evidence indicates. For example, in recent years mainstream economists have increasingly acknowledged the innovations of behavioral economics, which are now being absorbed into the mainstream, leading some to proclaim that behavioral economics simply *is* economics. Let's hope that sometime in the near future, we can say that ethics-and-economics simply *is* economics—which would certainly be a better economics, both practically and morally speaking.

Although the phrase “ethics and economics” would suggest a two-way exchange of ideas—which would be ideal—this has not been the pattern in the field, which, as described above, has dealt primarily with economists attempting to incorporate more elaborate conceptions of ethics into their work to better explain behavior or influence policy. For the most part, this handbook follows suit, with the chapters herein presenting various ways that ethics can contribute to economics. The volume is divided into two broad areas, foundations and applications, with a concluding chapter explaining what economics can contribute to ethics in turn, pointing the way to a true bilateral cooperation between the two fields in the future.

The first half of the book looks at foundational issues of ethics in economics. The first section surveys five significant ethical thinkers or schools of thought and explores their actual or potential influence on economics. Appropriately, the first chapter, by **Stefanie Haeffele** and **Virgil Henry Storr**, deals with Adam Smith, the moral philosopher cited at the beginning of this introduction, explaining how he established a more complete social science that emphasized sociality and sympathy, especially in markets (which are looked at in more depth in the second part of the volume). The second chapter, by **Jennifer A. Baker**, explores the contributions that a rich conception of virtue, based on ancient philosophers such as Aristotle and the Stoics (as well as Smith), can make to economics, presented in the context of business ethics—and a notorious character from the work of Charles Dickens. In the third chapter, **Mark D. White** surveys his work integrating the deontological ethics of Immanuel Kant into economic theory and practice, and argues that economists need to acknowledge concepts such as principles, rights, and duties in order to capture the richness of human choice, as well as the true impact of utilitarian policy that wrongfully harms some for the benefit of others. In the fourth chapter, **Jonathan B. Wight** argues for a pluralist approach to ethics in economics, writing that no single school of ethics can capture every relevant moral aspect of economic phenomena. He recommends combining and balancing the perspectives of virtue, duty, and welfare, and responds to criticisms of a pluralistic approach to ethics in economics. In the final chapter of this section, **Constanze Binder** and **Ingrid Robeyns** present the capabilities approach of modern philosophers Amartya Sen and Martha Nussbaum in relation to economics. As Binder and Robeyns argue, the capabilities approach balances the two concerns of freedom and well-being in a way that makes it particularly useful for the evaluation of institutions and economic systems, but not without limits, which they are also careful to identify.

The second section of the foundations half of the book looks at the evolutionary roots of morality in the context of economic theory and thought. The first chapter, by **Geoffrey M. Hodgson**, addresses the issue of sympathy versus self-interest highlighted by Adam Smith and others, but in the context of the work of Charles Darwin, who emphasized the same themes in his work, and also explored how language and deliberation were essential to the development of morality. On this basis, Hodgson explains how moral sentiments important to economic behavior may have evolved, and questions whether the assumption of other-regarding preferences is sufficient for economic models to explain altruistic behavior. In the second chapter, **Gerald Gaus** presents the work of Friedrich von Hayek, who used the concept of spontaneous order to offer an explanation of the evolution of morality that aligns with current science while avoiding the utopianism of many current political philosophers. Gaus challenges, however, Hayek's focus on group selection, arguing instead that a Smithean "invisible hand" can maintain social cooperation. In the final chapter of this section, **David C. Rose** surveys classical economic ideas about decision-making and introduces the visual metaphor of a bookshelf with books of different colors to represent choices of different degrees of social conformity. He then uses this to explain how the style of economic thinking that Deirdre McCloskey terms "Max-*U*" could have evolved in tandem with the development of the market economy.

The contributors to the third section of this half of the book explore various schools of thought within economics and how each incorporates ethics to better explain economic behavior. The first chapter, by **Sanjit Dhami** and **Ali al-Nowaihi**, surveys the work done by behavioral and experimental economists in studying ethical and unethical behavior—particularly lying—and argues that behavioral economics provides a more inclusive framework that results in more accurate and useful predictions of economic behavior, especially when ethical factors are present. The second chapter, by **John B. Davis**, examines the interrelationship between ethics and economics on a more methodological level, exploring the nature of the relationship between the disciplines using a complex systems approach. Davis argues that social economics, which assumes that market processes are embedded in social processes, offers a better context for dealing with ethics and economics than mainstream economics does. In the third chapter, **Peter J. Boettke** and **Kaitlyn Woltz** look at the connection between economics and ethics in the Austrian school of economics, with particular attention to the separation between ethical assumptions and economic analysis in the context of twentieth-century Austrian economists' support of free markets and capitalism (which are addressed further in the second half of the book). In the next chapter, **Ulrike Knobloch** explores the ethics of feminist economics, or what she calls *feminist economic ethics* (which also takes queer and postcolonial ethics into account). In this framework, she critically examines various topics such as the androcentric point of view of economics, the ethics and economics of gender norms, and the nature of ethics, justice, and provisioning, all from the point of view of care. Finally, the chapter by **Arlo Klammer** introduces a value-based approach to economics based on cultural economics, using a novel five-spheres model of the economy to explore the many ways that culture is essential to economics and emphasize that sense-making is an integral aspect of the economic process.

The second half of the book focuses on applications of ethics to important topics within economics, whether from a mainstream or expanded perspective. The first section is devoted to the ethics of commerce and markets, a topic mentioned in several of the chapters in the first half of the book. In his opening chapter, **James R. Otteson** discusses the ethical approaches of contributors to classical political economy, especially David Hume and Adam Smith, as well as their critics, including Jean-Jacques Rousseau and Karl Marx. He describes how Hume and Smith saw markets as not only moral but truly humane, helping to ensure for a broad prosperity, give special assistance to the poor, and enhance sociality in general. **Julian Reiss** takes a similar approach in his chapter, looking at the relationship between capitalism and democracy through the work of Alexis de Tocqueville, Joseph Schumpeter, and Karl Polanyi, as well as Marx and Hayek. Reiss then extends their perspectives into the current literature, in the end questioning which type of capitalism that exists today, if any, can co-exist with democracy. In the third chapter, **Joseph Heath** focuses on profit, pointing out its counterintuitive role as the motivation for competition, which ideally drives profit to zero. Because morality is normally understood to make people more cooperative, Heath asks if profit can truly be moral if it encourages competition, even as a means to overall beneficial end. Next, **Joakim Sandberg** examines the ethics of money and finance, starting with their existence

per se and then questioning the ethics of financial practices, suggesting ways to make them more honest or fair. He also considers arguments that agents within financial systems have social responsibilities beyond their own self-interest, a common question in discussions of commerce in general. Finally, **Michael S. McPherson** and **Debra Satz** look at labor markets, perhaps the most human aspect of commerce if not always the most humane. Recognizing that markets for labor have unique properties that distinguish them from other markets for inputs or output, they survey a number of perennial issues, including unemployment, the minimum wage, and the organization of work in a just society, as well as ethical implications of the rise of the gig economy.

The second section of this half of the book deals with the difficult problem of policy-making in an ethical fashion, both within welfare economics itself—especially in the context of risk—as well as in conjunction with fairness and equality. In the first chapter, **Matthew D. Adler** provides a comprehensive review of cost-benefit analysis and social welfare functions, the two most widely used approaches to welfare economics and policymaking, and makes a case that the social welfare function approach is preferable in the context of a number of important ethical principles that complement basic welfarism (including prioritarianism, which gives additional weight to the well-being of the worse-off). Next, **Marc Fleurbaey** focuses on risk and uncertainty, and examines several ethical questions with making social decisions involving risk to various parties, such as whether social decisions should be more or less risk-averse than those individuals make for themselves; how policymakers should respond to large but unlikely catastrophes compared to smaller but more frequent harms; and how these questions relate to egalitarianism, both *ex ante* and *ex post*. The third chapter, by **Luc Bovens**, looks at similar ethical issues of risk using examples such as drug allocations, charitable giving, breast-cancer screening, and Caesarian sections. He explains how each example has its own unique characteristics that make a utilitarian approach unpalatable, and questions the sufficiency of formal modeling alone to solve these difficult issues. Finally, **George F. DeMartino** looks at what he calls *econogenic harm*, harm done by economists in the process of trying to do good, which arises because of the uneven impact of economic interventions (especially those justified by cost-benefit analysis) and the epistemic problems that plague much work in economics. Because econogenic harm is inevitable to some degree, DeMartino proposes an “economy harm profile analysis” to help economists analyze it, reduce it, and identify opportunities to rectify it when possible.

The final section of this half of the book looks at four major applied fields of economics with significant ethical aspects. In the first chapter, **Daniel M. Hausman** tackles the fraught issues of health care and health insurance, explaining the difficulty of designing health policy and law to deal with asymmetric information, adverse selection, and moral hazard. He then complements the typical economic concern of efficiency with the equally important values of compassion, choice, efficiency, fairness, and solidarity, arguing that, in the end, health care markets must be both economically sound and morally defensible. Next, **Eyal Zamir** and **Barak Medina** consider deontological morality within the context of the economic approach to law (or *law and economics*) as a way to counterbalance and improve upon the standard utilitarian approach to the discipline.

Specifically, they propose using threshold deontology, which limits the enforcement of deontological rules when their cost reaches a particularly high level, as a way to integrate rules and standards within a legal decision-making framework without abandoning concern with outcomes altogether. In the third chapter, **David Schmidtz** investigates the concept of justice within discussions of ecology and the environment, arguing that an economic approach is integral to arriving at realistic solutions. After asking what it means for justice to be realistically ecological, he discusses the evolution of justice among early humans and how it makes compromise possible in a world of political conflict, especially with regard to environmental issues. In the last chapter of the section, **Brendan O’Flaherty** looks at the justification for civil rights in terms of both ethics and economics, explaining why they rightfully apply only to those subject to discrimination, and critically examines their effectiveness in light of their rationale, finishing with some suggestions for reform.

The volume concludes with a chapter by **John Broome** that turns the focus away from what ethics can contribute to economics, and points instead to lessons that economics can teach ethics, such as the right and (especially) wrong way to use the word “utility,” as well as the usefulness of mathematical methods, which he illustrates with an overview of the discussions of prioritarianism in both economics and philosophy (moral and political). More generally, Broome’s chapter suggests an avenue for the field of economics and ethics to be a truly interdisciplinary venture in which each part contributes to the other, possibly even leading to a new collaborative entity in which both ethics and economics work together toward their common goals of human prosperity and flourishing, however those might be defined or expressed.

REFERENCES

- Becker, Gary S. 1976. *The Economic Approach to Human Behavior*. Chicago, IL: University of Chicago Press.
- Bergstrom, Theodore C. 1999. “Systems of Benevolent Utility Functions.” *Journal of Public Economic Theory* 1: 71–100.
- Boettke, Peter J., and Christopher J. Coyne (eds.). 2015. *The Oxford Handbook of Austrian Economics*. Oxford: Oxford University Press.
- Bowles, Samuel, and Herbert Gintis. 2011. *A Cooperative Species: Human Reciprocity and Its Evolution*. Princeton, NJ: Princeton University Press.
- Davis, John B., and Wilfred Dolfsma (eds.). 2015. *The Elgar Companion to Social Economics*. 2nd edn. Cheltenham, UK: Edward Elgar.
- DeMartino, George F. 2011. *The Economist’s Oath: On the Need for the Content of Professional Economic Ethics*. Oxford: Oxford University Press.
- DeMartino, George F., and Deirdre N. McCloskey (eds.). 2016. *The Oxford Handbook of Professional Economic Ethics*. Oxford: Oxford University Press.
- Dhami, Sanjit. 2016. *The Foundations of Behavioral Economic Analysis*. Oxford: Oxford University Press.
- Dutt, Amitava, and Charles K. Wilber. 2010. *Economics and Ethics: An Introduction*. New York: Palgrave Macmillan.

- Etzioni, Amitai. 1987. "Toward a Kantian Socio-Economics." In Charles K. Wilber (ed.), *Economics, Ethics, and Public Policy* (Lanham, MD: Rowman & Littlefield), pp. 139–49.
- Etzioni, Amitai. 1988. *The Moral Dimension: Toward a New Economics*. New York: Free Press.
- Fehr, Ernst, and Klaus M. Schmidt. 1999. "A Theory of Fairness, Competition, and Cooperation." *Quarterly Journal of Economics* 114: 817–68.
- Fehr, Ernst, and Klaus M. Schmidt. 2006. "The Economics of Fairness, Reciprocity and Altruism—Experimental Evidence and New Theories." In Serge-Christophe Kolm and Jean Mercier Ythier (eds.), *Handbook of the Economics of Giving, Altruism and Reciprocity, Volume 1* (Dordrecht: Elsevier), pp. 615–91.
- Ferber, Marianne A., and Julie A. Nelson. 2003. *Feminist Economics Today: Beyond Economic Man*. 2nd edn. Chicago, IL: University of Chicago Press.
- Fischer, Liliann, Joe Hasell, J. Christopher Proctor, David Uwakwe, Zach Ward Perkins, Catriona Watson, et al. (eds.). 2017. *Rethinking Economics: An Introduction to Pluralist Economics*. London: Routledge.
- Hausman, Daniel, Michael McPherson, and Debra Satz. 2016. *Economic Analysis, Moral Philosophy, and Public Policy*. 3rd edn. Cambridge: Cambridge University Press.
- Margolis, Howard. 1984. *Selfishness, Altruism, and Rationality*. Chicago, IL: University of Chicago Press.
- McCloskey, Deirdre N. 2006. *The Bourgeois Virtues: Ethics in the Age of Commerce*. Chicago, IL: University of Chicago Press.
- McCloskey, Deirdre N. 2008. "Adam Smith, the Last of the Former Virtue Ethicists." *History of Political Economy* 40: 43–71.
- McCloskey, Deirdre N. 2010. *Bourgeois Dignity: Why Economics Can't Explain the Modern World*. Chicago, IL: University of Chicago Press.
- McCloskey, Deirdre N. 2016. *Bourgeois Equality: How Ideas, Not Capital or Institutions, Enriched the World*. Chicago, IL: University of Chicago Press.
- Peil, Jan, and Irene van Staveren (eds.). 2009. *Handbook of Economics and Ethics*. Cheltenham, UK: Edward Elgar.
- Sen, Amartya K. 1970. "The Impossibility of a Paretian Liberal." *Journal of Political Economy* 78: 152–57.
- Sen, Amartya K. 1977. "Rational Fools: Fools: A Critique of the Behavioral Foundations of Economic Theory." *Philosophy & Public Affairs* 6: 317–44.
- Wight, Jonathan B. 2015. *Ethics in Economics: An Introduction to Moral Frameworks*. Stanford, CA: Stanford University Press.

PART I

FOUNDATIONS

A. ETHICAL THEORIES IN ECONOMICS

CHAPTER 1

ADAM SMITH AND THE STUDY OF ETHICS IN A COMMERCIAL SOCIETY

STEFANIE HAEFFELE AND
VIRGIL HENRY STORR

1.1 INTRODUCTION

ALTHOUGH economics began as a branch of moral philosophy, the study of markets and the study of ethics are now mostly pursued separately as distinct phenomena.¹ Within neoclassical economics, for instance, the emphasis is typically on the rationality of market actors and how their behavior is shaped by incentives. The individuals acting within the market, when analyzed this way, appear to be cold calculators of benefits and costs—the *homo economicus* of neoclassical economic modeling. Their preferences are simplified into aggregable utility functions and their motivations reduced to external incentive structures. The market is viewed as an amoral space where rational actors in certain settings interact in ways that advance social progress (that is, by the invisible hand) and in other settings fail to achieve positive social outcomes (or market failure). Arguably, for most economists, the moral dispositions of market actors and the morality of markets are beyond, or at most tangential to, their scholarly focus.

Contemporary economists, thus, face a challenge of how to incorporate the ethical orientations of individuals into their analysis.² The strictly rational *homo economicus* can have no ethical considerations without impugning his rationality. Yet, in daily life, economics and ethics (as well as politics, charity, and other aspects of human interaction)

¹ Sen (1977) argues that this distinction was solidified when Edgeworth published *Mathematical Physics* in 1881.

² For example, Sen (1977) suggests that focusing on social commitment is a way to incorporate ethics into economics.

are interconnected.³ The same person attends church on Sunday, goes to work at a firm on Monday, votes on Tuesday, purchases groceries on Wednesday, and volunteers at a homeless shelter on Thursday. It is unsurprising that in the real world, one's ethics, goals, and experiences shape their actions in all these areas (see, for instance, Granovetter 1985; Macneil 1986; Storr 2008).

Behavioral and experimental economists, however, have not only pointed out some of the flaws within the neoclassical conception of rationality, but have also shown how people incorporate ethics into market transactions.⁴ Indeed, the hopelessly flawed and often irrational beings—that is, the imperfect actors of behavioral economics—who often fail to achieve their personal goals, are ethical creatures who attempt to but do not always live up to their ethical commitments. For example, researchers have found that individuals value trustworthiness in their trading partners (Kumar 1996), punish dishonesty and greed in trade (Akerlof 1970; Adler 1992), and prefer firms that have high employee satisfaction, high customer loyalty, and overall good reputations (Loveman 1998; Roberts and Dowling 2002). Furthermore, scholars have argued that markets provide opportunities for social interaction and moral learning (McCloskey 2006; Novak 1984; Paganelli 2010; Storr 2008, 2010). Arguably, despite these efforts, the question of how best to incorporate ethical concerns into economic analysis, without having to abandon traditional insights about the nature and functioning of a market system, remains unsettled.

A fuller examination of the market requires understanding and incorporating ethical considerations and dispositions into the study of markets. Contemporary economists might benefit from looking to Adam Smith as an exemplar for how the study of political economy can incorporate ethics and economics. Smith is known for two seminal works: *The Inquiry into the Nature and Causes of Wealth of Nations* (1776, henceforth WN), which studies markets, and *The Theory of Moral Sentiments* (1759, henceforth TMS), which studies ethics. While these two distinct topics may be seen as contradictory (known as “Das Adam Smith Problem”), scholars have recognized that they are inextricably linked.⁵

In WN, Smith argues that not only should we not rely on or expect benevolence from others in order to survive and thrive, we thrive precisely because others are seeking to improve their own lot in life (Berry 1990; Muller 1993; Rasmussen 2008; Rothschild 2001; Werhane 1991). As Smith famously stated, “It is not from the benevolence of the butcher, the brewer, or the baker, that we expect our dinner, but from their regard to their own interest. We address ourselves, not to their humanity but to their self-love, and never talk to them of our own necessities but of their advantages” (26–7).

This self-interest, or self-love as Smith calls it, which motivates individuals to pursue market exchange, is also constrained by our moral sentiments (Storr 2018). In TMS,

³ See Rothschild (2001) for Smith's view of this complex and interconnected nature of social life.

⁴ See Ashraf, Camerer, and Loewenstein (2005) for a discussion of Smith as an early behavioral economist.

⁵ For an overview and assessment of the debate on Das Adam Smith Problem, see Montes (2003) and Teichgraber (1981). For arguments supporting the linkages between the two works, see for instance Macfie (1959), Smith (1998), Coker (1990), Rothschild (2001), Otteson (2002), Dwyer (2005), Fleischacker (2004), Paganelli (2008), Rasmussen (2008), Herzog (2016), Pitts (2017), and Storr (2018).

Smith argues that we seek to share our passions and reactions with others in order to feel understood and to legitimate our feelings and actions (that is, we seek fellow-feeling) (Den Uyl 1991; Griswold 1999). Smith observes that “nothing pleases us more than to observe in other men a fellow-feeling with all the emotions of our own breast” (13) and “to approve of another man’s opinions is to adopt those opinions, and to adopt them is to approve of them” (17). Our feelings and actions are legitimate when we receive approval and praise from others. Yet, we do not just want to receive praise; we also want to be truly praiseworthy. As Smith states, “Nature, accordingly, has endowed him, not only with a desire of being approved of, but with a desire of being what ought to be approved of; or of being what he himself approves of in other men” (117). In order to assess praiseworthiness of ourselves and others, we utilize an *impartial spectator*, or man in the breast, as an unbiased judge. Smith argues that in order to judge mankind, individuals appeal “to a much higher tribunal, to the tribunal of their own consciences, to that of the supposed impartial and well-informed spectator, to that of the man within the breast, the great judge and arbiter of their conduct” (130).

Taken together, Smith suggests that we seek to improve our lot in life but also to be understood and approved of by our peers and the impartial spectator. These dual forces shape our social interactions. In order to understand the world around us, both aspects—economics and ethics—should be explored in our studies. Specifically, as students of commercial society, economists ought to adopt this Smithian strategy and tackle questions about the interconnection between economics and ethics.

Furthermore, we argue that this Smithian approach to political economy has implications about the market and society. First, markets do not need to be defended by claiming their amorality, but instead can be a space for moral interaction and development (see also Storr 2018). Second, a fuller understanding of individuals as social creatures allows us to see how our lives are shaped by markets, along with civil society, politics, and more.

The chapter proceeds as follows. Section 1.2 examines Smith’s theory of markets and his theory of ethics is explored in Section 1.3. Section 1.4 provides a synthesis of these theories into a Smithian theory of markets that focuses on ethics, and Section 1.5 concludes.

1.2 ADAM SMITH’S THEORY OF MARKETS

In WN, Smith shows how markets, through specialization and the division of labor, have led to incredible social progress. Indeed, he opens the book with: “The greatest improvement in the productive powers of labour, and the greater part of the skill, dexterity, and judgement with which it is any where directed, or applied, seem to have been the effects of the division of labour” (1). With the division of labor, jobs are divided into sequences of specific tasks. The making of a pin, woolen coat, or pencil requires the efforts of thousands: from collecting the raw materials, to constructing the individual

components, to shipping and selling the final products.⁶ This specialization does not just occur within manufacturing but throughout the economy. For instance, lawyers often specialize in practicing specific types of law, such as corporate, family, or environmental law.

The division of labor allows for workers to produce more than they ever could on their own, by parsing out tasks and gaining skills in particular areas. As Smith states:

This great increase of the quantity of work, which, in consequence of the division of labour, the same number of people are capable of performing, is owing to three different circumstances; first, to the increase of dexterity in every particular workman; secondly, to the saving of the time which is commonly lost in passing from one species of work to another; and lastly, to the invention of a great number of machines which facilitate and abridge labour, and enable one man to do the work of many.

(WN: 17)

While improved dexterity and reduced switching costs allow workers to produce more during work hours, labor-saving inventions are an attempt to reduce or shift their workload. By finding new and creative ways to fulfill their tasks, such as through automation, workers can simplify their tasks or take on new responsibilities. As Smith explains:

It is naturally to be expected, therefore, that some one or other of those who are employed in each particular branch of labour should soon find out easier and readier methods of performing their own particular work, wherever the nature of it admits of such improvement. (WN: 20)

This diverse and intricate nature of the market is not only found in the cities where we live, but stretches across nations. We shop for groceries at our local supermarket which employs dozens of local residents and whose inventory comes from all over the world. When we purchase coffee, spices, and produce, they often come from communities we have never travelled to and from people we will never meet. The wide array of goods and services available shows the global extent of the market in modern developed nations. Indeed, Smith notes that progress is constrained by the extent of the market:

As it is the power of exchanging that gives occasion to the division of labour, so the extent of this division must always be limited by the extent of that power, or, in other words, by the extent of the market. When the market is very small, no person can have any encouragement to dedicate himself entirely to one employment, for want of the power to exchange all that surplus part of the produce of his own labour, which is over and above his own consumption, for such parts of the produce of other men's labour as he has occasion for. (WN: 31)

The goods, services, and social progress that comes with the market are available to all in society who can work for a wage or who otherwise have funds to purchase them, no

⁶ See Read (1958) for an in-depth discussion of the division of labor required to produce a single pencil.

matter how kind or mean, generous or miserly, benevolent or selfish. As Smith notes, “without the assistance and co-operation of many thousands, the very meanest person in a civilized country could not be provided, even according to, what we very falsely imagine, the easy and simple manner in which he is commonly accommodated” (WN: 23). Members of society benefit from the progress of the market, not because of the benevolence of their neighbors, but because social progress is possible when everyone aims to improve their own lot in life (Muller 1993; Otteson 2002; Rasmussen 2008; Rothschild 2001). In other words, pursuing our individual self-interest is incentive-compatible with social progress.

Additionally, in order to achieve our own goals, we must exchange with, interact with, and rely on others in society. This cooperation is another incentive-compatible aspect of the market; it is not from benevolence but self-interest that people exchange with and cooperate with one another. As Smith observes,

man has almost constant occasion for the help of his brethren, and it is in vain for him to expect it from their benevolence only. He will be more likely to prevail if he can interest their self-love in his favour, and shew them that it is for their own advantage to do for him what he requires of them. Whoever offers to another a bargain of any kind, proposes to do this: Give me that which I want, and you shall have this which you want, is the meaning of every such offer; and it is in this manner that we obtain from one another the far greater part of those good offices which we stand in need of. It is not from the benevolence of the butcher, the brewer, or the baker, that we expect our dinner, but from their regard to their own interest. We address ourselves, not to their humanity but to their self-love, and never talk to them of our own necessities but of their advantages. (WN: 26–27)

Smith saw this process as resulting in more than just the culmination of individuals pursuing their self-interest and achieving personal success (Herzog 2016; Rasmussen 2006; Rothschild 2001). Social progress impacts all members of society, as even the poorest amongst us have access to and can benefit from the knowledge and innovations that were once luxuries only available to kings and queens.⁷ Further, knowledge and innovation spreads throughout the market, leading to new adoptions and adaptations. Smith argued that this spreading of knowledge through the market would lead to further progress:

nothing seems more likely to establish this equality of force than that mutual communication of knowledge and of all sorts of improvements which an extensive commerce from all countries to all countries naturally, or rather necessarily, carries along with it. (WN: 627)

⁷ “Compared, indeed, with the more extravagant luxury of the great, his accommodation must no doubt appear extremely simple and easy; and yet it may be true, perhaps, that the accommodation of an European prince does not always so much exceed that of an industrious and frugal peasant, as the accommodation of the latter exceeds that of many an African king, the absolute master of the lives and liberties of ten thousand naked savages” (WN: 23–4).

While the market is not perfect—it does not lead to benefits and progress for all individuals at all times, as we discuss in Section 1.4—the trend of social interaction within the market is toward social progress. This is possible, Smith advances, through the division of labor and the realization that social cooperation is not dependent on benevolence. Even when our neighbors are mean, miserly, or selfish, Smith argues in WN, we can still cooperate and prosper together. Next, we turn to ethics and Smith’s understanding of how individuals are not just self-interested but also social, ethical beings.

1.3 ADAM SMITH’S THEORY OF ETHICS

Smith’s theory of ethics is in the tradition of the Enlightenment philosophers of his time, building off of but distinguishing itself from the works of his teacher, Francis Hutcheson, and his friend, David Hume (Griswold 1999; Hope 1989; Raphael 1972–1973; Rasmussen 2008).⁸ As McCloskey (2008) explains, Smith focused on five virtues from pagan, stoic, and Christian traditions: courage, temperance, justice, prudence, and love. While his focus on virtues was consistent with virtue ethics, Smith also advanced a notion of ethics that was practical (that is, useful in engaging in social life) and not simply as a higher order good of its own. Rather than seeking virtue for its own sake, for Smith individuals seek virtue because of their social nature (Schliesser 2006).

As Smith advances in WN, we are inherently social beings, dependent on our neighbors as well as strangers in order to go about our daily lives. In addition to needing others to obtain goods and services, Smith argues in TMS that we also seek the understanding and approval of others, and when we see others in pain, we often think about how it must feel to be in their situation and are often pained as well: “That we often derive sorrow from the sorrow of others, is a matter of fact too obvious to require any instances to prove it” (TMS: 9). We imagine ourselves in their shoes and attempt to understand the situation through their eyes; we attempt to make their situation our own. According to Smith,

By the imagination we place ourselves in his situation, we conceive ourselves enduring all the same torments, we enter as it were into his body, and become in some measure the same person with him, and thence form some idea of his sensations, and even feel something which, though weaker in degree, is not altogether unlike them. (ibid.)

Smith refers to this ability to deploy our moral imagination to place ourselves in others’ circumstances as *sympathy* or *fellow-feeling* (Den Uyl 1991; Griswold 1999).

Just as we can engage in sympathy, we also seek it from others. As Smith notes, “nothing pleases us more than to observe in other men a fellow-feeling with all the

⁸ Both Smith and Hume focus on but have different approaches to sympathy (Levy and Peart 2004; Morrow 1923) and justice (Raphael 1972–1973). For contemporary research on Adam Smith and his influences, see Montes and Schliesser (2006) and Thomas (2015).

emotions of our own breast; nor are we ever so much shocked as by the appearance of the contrary” (TMS: 13). Others sharing in our triumphs and downfalls “enlivens joy and alleviates grief. It enlivens joy by presenting another source of satisfaction; and it alleviates grief by insinuating into the heart almost the only agreeable sensation which it is at that time capable of receiving” (TMS: 14). Furthermore, to imagine yourself in such circumstances and to share in someone’s joy or sorrow, is to legitimize their reactions and feelings:

To approve of another man’s opinions is to adopt those opinions, and to adopt them is to approve of them. If the same arguments which convince you, convince me likewise, I necessarily approve of your conviction; and if they do not, I necessarily disapprove of it. (TMS: 17)

This type of social interaction provides feedback about the relative appropriateness of our feelings and actions. In other words, sympathy acts as a mirror, allowing us to see how others perceive us.⁹ Smith notes the importance of society and community in living a well-balanced life, stating,

Society and conversation, therefore, are the most powerful remedies for restoring the mind to its tranquility, if, at any time, it has unfortunately lost it; as well as the best preservatives of that equal and happy temper, which is so necessary to self-satisfaction and enjoyment. (TMS: 23)

It is not enough, however, to receive praise from others for our passions and actions. We also seek to be praiseworthy, to be virtuous (Den Uyl 1991; Griswold 1999). Smith notes that nature “has endowed [man], not only with a desire of being approved of, but with a desire of being what ought to be approved of; or of being what he himself approves of in other men” (TMS: 117).¹⁰ To be praiseworthy is to pursue appropriate actions and feel appropriate passions in regard to those circumstances. And while we seek to be viewed as praiseworthy by our peers, we also know that our biases can alter our judgements of ourselves and others. Therefore, Smith argues, we employ the assessment of an imaginary impartial spectator.

This spectator, stripped of bias and endowed with knowledge and experiences, can judge our actions and passions from an uncompromised position. Smith distinguishes the self from the impartial spectator as follows:

When I endeavor to examine my own conduct, when I endeavor to pass sentence upon it, and either to approve or condemn it, it is evident that, in all such cases, I divide myself, as it were, into two persons; and that I, the examiner and judge, represent a

⁹ For the importance of society as a social mirror for identifying and legitimizing virtues, see Berry (1990).

¹⁰ In TMS: 126–7, Smith notes that the desire for praise and the desire to be praiseworthy are related but separate desires. The first is to be favorable and the second is to be proper and virtuous, though both are driven by a desire to avoid being blamed or blameworthy.

different character from that other I, the person whose conduct is examined into and judged of. The first is the spectator, whose sentiments with regard to my own conduct I endeavour to enter into, by placing myself in his situation, and by considering how it would appear to me, when seen from that particular point of view. The second is the agent, the person whom I properly call myself, and of whose conduct, under the character as a spectator, I was endeavouring to form some opinion. The first is the judge; the second the person judged of. (TMS: 113)

The impartial spectator aids our assessment of our own actions and those of others, providing a consistent metric for judgement.¹¹ Additionally, when the impartial spectator approves (or disapproves) of our actions, it is likely that our peers will as well. As Smith notes:

That degree of self-estimation, therefore, which contributes most to the happiness and contentment of the person himself, seems likewise most agreeable to the impartial spectator. The man who esteems himself as he ought, and no more than he ought, seldom fails to obtain from other people all the esteem that he himself thinks due. (TMS: 261)

Thus, employing the impartial spectator not only provides us with an unbiased judgement of our actions and feelings, but also aids us in being approved of by our community.¹²

1.4 ADAM SMITH ON ETHICS AND MARKETS

Viewed together, WN and TMS provide a complex view of social interaction within markets (Fleischacker 2004; Macfie 1959; Otteson 2002; Pitts 2017; Rasmussen 2008).¹³ First, individuals seek to improve their lives and, by pursuing this self-interest within

¹¹ Smith further explains impartial spectator compared to the individual as follows: “though man has...been rendered the immediate judge of mankind, he has been rendered so only in the first instance; and an appeal lies from his sentence to a much higher tribunal, to the tribunal of their own consciences, to that of the supposed impartial and well-informed spectator, to that of the man within the breast, the great judge and arbiter of their conduct. The jurisdictions of those two tribunals are founded upon principles which, though in some respects resembling and akin, are, however, in reality different and distinct. The jurisdiction of the man without is founded altogether in the desire of actual praise, and in the aversion to actual blame. The jurisdiction of the man within is founded altogether in the desire of praiseworthiness, and in the aversion to blameworthiness” (TMS: 130–1).

¹² While Smith views the impartial spectator as universal—everyone has access to the man in their breast—its knowledge and experiences are grounded by culture and community (that is, the spectator’s cultural lenses) (Forman-Barzilai 2005; Gibbard 1992; Storr 2018; Storr and John 2015). As we interact with others, we expand our experiences as well as those of the impartial spectator.

¹³ See also Thomas (2015), an edited volume on sympathy and selfishness, on Smith’s combination of economics and ethics.

the market, they promote social progress. Second, individuals seek to be understood and legitimized by their peers and the impartial spectator. These dual forces shape our social interactions. And, this approach to studying social interaction illuminates a fuller picture of commercial society. A theory of markets that focuses on ethics should recognize that markets are the source of the wealth of nations. A theory of markets that focuses on ethics should also recognize (1) the limits of benevolence, (2) the potential for corruption, and (3) the possibility of moral development.

1.4.1 The Limits of Benevolence

For Smith, *benevolence* (or as he calls elsewhere *beneficence*) describes the attention, care, and generosity that we extend to others. Although we are “more sensibly” concerned with our own pleasures and pains than those of others, we are not naturally indifferent to the well-being of others. In fact, we want to enliven the joys of others and alleviate their suffering, even at our own expense. To say, though, that we are naturally beneficent and to celebrate altruism as a virtue is not to suggest that we are naturally, or that we ought to be, self-neglecting or self-sacrificing. According to Smith, the perfectly virtuous man “acts according to the rules . . . of *proper* benevolence” (TMS: 237, emphasis added). The rules of “*proper* benevolence,” however, do not require that we elevate the concerns of others above concern for ourselves. As such, there are naturally and appropriately limits to our generosity.

Prudence is not the only check on our beneficence. Our passions sometimes prevent us from being altruistic; our natural parochialism focuses our attention and affections on those most proximate to us, and the less intimate our connections with someone the more difficult it is to know that they need help and how to help them (Forman-Barzilai 2005, 2010; Hanley 2008; Otteson 2002; Paganelli 2010; Storr 2018). As Smith acknowledges, knowing the rules of proper benevolence is not enough to guarantee that someone acts in accordance with them. “His own passions,” he explains, “are very apt to mislead him; sometimes to drive him and sometimes to seduce him to violate all the rules which he himself, in all his sober and cool hours, approves of” (TMS: 237). Our passions could lead us to give more generously than we are able to when we are confronted with another person’s sad circumstances or might lead us to give less generously than we might otherwise when we are fearful that helping them might harm us. Smith suggests that self-command could check our passions and ensure that we act as we ought to act, but also acknowledges that “our sensibility . . . both to our own injuries and to our own misfortunes, though generally too strong, may likewise be too weak” (TMS: 244).

In addition to our passions sometimes leading us to be more or less altruistic than we could and should be, we often direct our altruism toward those we know best. That we concern ourselves first with ourselves, then with our families, then with our nations, and only then with all of mankind, should not be viewed as a bug in our moral systems but a feature. While self-love should not lead us to harm a stranger in order to help ourselves or our familiars, it is an appropriate exercise of our benevolence to

focus first on our intimates.¹⁴ Smith's thought experiment in regarding how we might react to learning of a devastating earthquake in some remote part of the world is telling (TMS: 136). Imagine, Smith suggests, how someone with no connection to that part of the world would be affected by learning of the disaster. According to him, it is likely that he will reflect on the horrors that befell them and maybe imagine how it might affect those more proximate to himself, but that his sleep that night is unlikely to be disturbed. If he were threatened with the loss of his little finger tomorrow, however, he would not be able to sleep tonight. It is also likely, though Smith does not extend the thought experiment in this direction, that potential calamity to the man's family or friends or nation would likewise trigger a restless repose. Again, in Smith's view, this is not a moral indictment but, rather, a moral fact. The implication of this, however, is clear: we cannot count on the care of those not proximate to us.

Even if we could expect generosity from those we do not have a connection with, we still might not be able to rely on it. The less intimate another person is to us, the more difficult it is to know that they need help and how best to help them. Stated another way, there is an *ethical knowledge problem* that limits our ability to do good as we move away from the people we know best. As Smith explains,

The administration of the great system of the universe, however, the care of the universal happiness of all rational and sensible beings, is the business of God and not of man. To man is allotted a much humbler department, but one more suitable to the weakness of his powers, and the narrowness of his comprehension: the care of his own happiness, of that of his family, his friends, his country. (TMS: 237)

As we move to less and less proximate "circles of sympathy," the weaker the claims on our beneficence, and also the weaker our capacity to appropriately direct our beneficence (Forman-Barzilai 2010).

Arguably, it is at least in part because our hopes for benevolence are necessarily limited that we appeal to self-love and not to benevolence when we seek the aid of others in the market. Again, as Smith explains, "man has almost constant occasion for the help of his brethren, and it is in vain for him to expect it from their benevolence only" (WN: 26). Even in the setting of charity, we cannot completely rely on others' benevolence. Specifically, Smith observes that,

Nobody but a beggar chuses to depend chiefly upon the benevolence of his fellow-citizens. *Even a beggar does not depend upon it entirely.* The charity of well-disposed people, indeed, supplies him with the whole fund of his subsistence. But though this principle ultimately provides him with all the necessities of life which he has occasion for, it neither does nor can provide him with them as he has occasion for them. The greater part of his occasional wants are supplied in the same manner as those of other people, by treaty, by barter, and by purchase. (WN: 27, emphasis added)

¹⁴ For a discussion on how self-interest is not selfish, see Griswold (1999).

The money and goods we give away to others are often earned or purchased in the market. Moreover, charitable organizations often rely on donations that can be traced back to profits or wages, often pay their employees, and often purchase the goods and services they utilize in their operations.

There are limits to how much benevolence we should expect from others. Altruism, as Robertson (1956) suggests, is the ultimate scarce resource and the market works in part because it allows us to economize on love. This is not to suggest that the market crowds out benevolence but, instead, suggests that the market can fill in for benevolence where we need one another but our benevolence is necessarily limited.

1.4.2 The Potential for Corruption

Although Smith was certainly a champion of commercial society, he recognized that markets presented us with some ethical conundrums (Force 2003; Hill 2006). For instance, he repeatedly complained about the “natural selfishness and rapacity” of the rich (TMS: 184; see also Force 2003 and Rosenberg 1979). In both WN and TMS, it is clear that Smith believed that the market (1) depended on the moral degradation, of at least some of us, and (2) often resulted in collusion, exploitation, alienation, and the corruption of our morals.

Take, as an example, the poor man’s son that Smith discusses in TMS. The poor man’s son grows up poor, observing the conveniences and luxuries of the rich. He connects their commercial successes with luxury, leisure time, and happiness and, therefore, sets out to become rich in order to obtain happiness.¹⁵ As Smith notes,

He thinks if he had attained all these, he would sit still contentedly, and be quiet, enjoying himself in the thought of the happiness and tranquillity of his situation It appears in his fancy like the life of some superior rank of beings, and, in order to arrive at it, he devotes himself for ever to the pursuit of wealth and greatness.
(TMS, 181)

So, he studies and works day and night for years in his ambitious quest for wealth.

Yet, when he is old and dying, the poor man’s son realizes that he did not achieve the leisure time and happiness he was after (Hont and Ignatieff 1983). As Smith stated,

It is then, in the last dregs of life, his body wasted with toil and diseases, his mind galled and ruffled by the memory of a thousand injuries and disappointments which he imagines he has met with from the injustice of his enemies, or from the perfidy and ingratitude of his friends, that he begins at last to find that wealth and greatness are mere trinkets of frivolous utility. (TMS: 181)

¹⁵ For an examination of self-deceit in Smith, see Fleischacker (2011). For an examination of happiness in Smith, see Rasmussen (2006).

While the poor man's son is led astray by the lure of riches, society benefits from his quest for wealth (Fleischacker 2004). According to Smith, "it is well that nature imposes upon us in this manner. It is this deception which rouses and keeps in continual motion the industry of mankind" (TMS: 183). Thus, for Smith, this corrupted view of riches is actually needed in order for society to make major advancements.¹⁶

Smith recognized that there is also a dark side of the market in its potential to promote collusion as well as worker exploitation and alienation. With the potential for human beings to exploit, alienate, coerce, and harm each other and themselves, Smith also recognized the need for justice. In *Lectures in Jurisprudence* (1982, henceforth LJ), Smith sees justice as "the security from injury" (398). Individuals, according to him, may be injured as a person through injuries of their body, reputation, or property. Thus, justice requires not only protecting ourselves and others from bodily and emotional harm but also protecting property rights (399).¹⁷ Smith argues that justice "is the foundation of civil government" (398), but also that people have a personal obligation to protect others from harm. In TMS, Smith states that

Wherever prudence does not direct, wherever justice does not permit, the attempt to change our situation, the man who does attempt it, plays at the most unequal of all games of hazard, and stakes every thing against scarce any thing. (149–50)

Take first, the incentive for industry leaders to get together to attempt to raise prices, decrease quality, limit supply, or find other ways to gain profit in ways their customers would not prefer (Rosenberg 1979). According to Smith, "People of the same trade seldom meet together, even for merriment and diversion, but the conversation ends in a conspiracy against the public, or in some contrivance to raise prices" (WN: 145). Although justice allows for us to compete, justice does not allow us to cheat to win. As Smith explains,

in the race for wealth, and honours, and preferments, he may run as hard as he can, and strain every nerve and every muscle, in order to outstrip all his competitors. But if he should juggle, or throw down any of them, the indulgence of the spectators is entirely at an end. It is a violation of fair play, which they cannot admit of. (TMS: 83)

Thus, the collusion that Smith suggests is not only inevitable but also immoral. That standards of living increase despite these efforts is because the pressure to collude is checked both by the sense of justice possessed by some industry leaders who refuse to collude in order to disadvantage the public as well as the propensity of members of a cartel to work to undermine the cartel.

Likewise, Smith worried that the power dynamic between employers and employees could potentially lead to worker exploitation and alienation. Employers attempt to pay

¹⁶ A modern-day example may be the ambition and personality of Steve Jobs, whose technological vision led to dozens of world-changing innovations.

¹⁷ See Raphael (1972–1973) for a discussion of ethics and justice in Hutcheson, Hume, and Smith.

the lowest possible wage that will sustain their workforce, and because they often have more resources, they hold more power when negotiating with their staff.¹⁸ As Smith notes, “In the long-run the workman may be as necessary to his master as his master is to him; but the necessity is not so immediate” (WN: 84). Employees seeking better wages or working conditions often must rally together in order to shift the power-dynamic with their employers. As he writes,

In order to bring the point to a speedy decision, [employees] have always recourse to the loudest clamour, and sometimes to the most shocking violence and outrage. They are desperate, and act with the folly and extravagance of desperate men who must either starve, or frighten their masters into an immediate compliance with their demands. (WN: 84–5)

The necessary recourse to “folly and extravagance” is evidence of their powerlessness.

When there is competition and progress, however, employees may have the advantage. As Smith argues,

When in any country the demand for those who live by wages; labourers, journeymen, servants of every kind, is continually increasing; when every year furnishes employment for a greater number than had been employed the year before, the workmen have no occasion to combine in order to raise their wages. The scarcity of hands occasions a competition among masters, who bid against one another, in order to get workmen, and thus voluntarily break through the natural combination of masters not to raise wages. (WN: 86)

As wealth increases, the demand for laborers increases, as do their wages. Moreover, as the economy advances, the more specialized the skills needed to perform particular jobs, the smaller the pool of potential laborers who possess those specific skills, and the higher the wages for that job will be. Further, Smith observes:

It is not the actual greatness of national wealth, but its continual increase, which occasions a rise in the wages of labour. It is not, accordingly, in the richest countries, but in the most thriving, or in those which are growing rich the fastest, that the wages of labour are highest. (WN: 87)

Said another way, the prosperity of workers is connected to social progress: when the economy is progressing, wages increase; when the economy is stagnating, wages stagnate; when the economy is in decline, wages decline.

Although Smith seems to conclude that exploitation is possible and likely, especially when the economy is not performing well, he seems to believe that alienation is

¹⁸ Scholars since Marx (1867) have taken up the critiques of the markets found in Smith. For a discussion of the potential of exploitation in the market, and whether the market or government is best suited to remedy such exploitation, see Storr and Haeffele (2016).

inevitable in a modern commercial society characterized by the division of labor (Griswold 1999; West 1975). According to him:

In the progress of the division of labour, the employment of the far greater part of those who live by labour . . . comes to be confined to a few very simple operations . . . The man whose whole life is spent in performing a few simple operations, of which the effects too are, perhaps, always the same, or very nearly the same, has no occasion to exert his understanding, or to exercise his invention in finding out expedients for removing difficulties which never occur. He naturally loses, therefore, the habit of such exertion, and generally becomes as stupid and ignorant as it is possible for a human creature to become . . . His dexterity at his own particular trade seems, in this manner, to be acquired at the expence of his intellectual, social, and martial virtues. (WN: 781–2)

This is a substantial critique. According to Smith, the division of labor narrows employees' tasks to such an extreme that they become incapable of interacting with their fellow members in society (Fleischacker 2004; Griswold 1999; Muller 1993; Pack 1991; West 1975; Winch 1992; Viner 1928). In response, Smith recommends public education as a way to diversify the activities and opportunities available to workers. Additionally, his own discussion of the division of labor offers an additional remedy: the incentive for labor-reducing innovations. Nonetheless, the moral risk associated with engaging in market activity is profound.

1.4.3 The Possibility of Moral Development

While a Smithian theory of markets that focuses on ethics helps highlight the limitations and undesirable aspects of the market, it also illuminates how the market can be a setting for moral development. Storr (2018) argues that the market will tend to reward virtuous behavior and punish vice, and as such should be viewed as a moral teacher. Specially, he finds that markets (1) reward our concern for others, (2) reward good behavior and punishes dishonest behavior, and (3) expand our moral communities.

Recall that Smith said that “it is not from the benevolence of the butcher, the brewer, or the baker, that we expect our dinner, but from their regard to their own interest” (WN: 26). Although we will not succeed in the market if we rely on others being particularly concerned about us, if we hope to succeed in the market we have to constantly concern ourselves with the desires and well-being of others. The successful entrepreneur is the one who figures out what goods and services potential buyers want them to sell, what kinds of arrangements are likely to attract key suppliers, and what kinds of commitments are likely to retain the best employees. The successful entrepreneur is the one who is always focused on how others will benefit from the deals that she is proposing. It is important to note that this concern for others extends to a concern about how those that we deal with in the market wish to be treated and the kind of trading partners we wish to interact. Stated another way, in the market there are rewards to be had for worrying about others, catering to their desires, and treating them as they would

want to be treated. To the extent that people wish to be treated decently, we should expect successful market actors to discover what this means and to treat others accordingly.

The market will, thus, tend to reward good behavior and also discipline or punish bad actors. As such, we should expect that successful market actors will tend to keep their commitments and to avoid even appearing to be dishonest. As Smith observed,

Whenever commerce is introduced into any country, probity and punctuality always accompany it. . . . It is. . . . reduceable to self interest, that general principle which regulates the actions of every man, and which leads men to act in a certain manner from views of advantage. (LJ: 538)

Furthermore:

A dealer is afraid of losing his character, and is scrupulous in observing every engagement. When a person makes perhaps 20 contracts in a day, he cannot gain so much by endeavoring to impose on his neighbors, as the very appearance of a cheat would make him lose. (ibid.)

Individuals will pay a premium to deal with honest trading partners and will tend to avoid trading partners who are dishonest (Adler 1992; Kumar 1996; Loveman 1998). As such, in markets, where repeated dealings are not only desirable but essential, and a bad reputation would be fatal to a business' prospects, integrity is profitable. The typical merchant interacts with hundreds of customers each day and could not afford for any of them to view him as a bad actor. Moreover, in commercial societies fidelity becomes commonplace. "When the greater part of people are merchants," Smith notes, "they always bring probity and punctuality into fashion, and these therefore become the principal virtues of a commercial nation" (LJ: 539).

The tendency for the market to reward caring for others may just be an incentive to act, or pretend, as if one cares for others. Say, for instance, a shopkeeper who realizes he is losing exchange opportunities because of his dishonest behavior may begin to act as if he were a kind and honest man in order to garner more business. He is persuaded to behave in an appropriate way, yet his actions may be insincere. While it is socially beneficial that he at least pretends to behave in this way, he may not actually become more virtuous. However, in order to maintain this status in his community and succeed in his business long term, he must continue to behave in this manner. Over time, it is likely that his once intentional actions will become instinctive and more genuine, and eventually result in actual moral development. Stated another way, a truly dishonest and conniving person is unlikely to convincingly pretend to be reputable for an extended period of time without being impacted by some sort of moral development.

Not only does the market reward us for caring about others, and rewards virtue while punishing vice, our market dealings lead us to care about the opinions and well-being of more people than we otherwise would. As Hayek explains:

That interdependence of all men, which is now in everybody's mouth and which tends to make all mankind One World, not only is the effect of the market order but

could not have been brought about by any other means. What today connects the life of any European or American with what happens in Australia, Japan or Zaire are repercussions transmitted by the network of market relations.... The benefits from the knowledge which others possess, including all the advances of science, reach us through channels provided and directed by the market mechanism. Even the degree to which we can participate in the aesthetic or moral strivings of men in other parts of the world we owe to the economic nexus. (1978: 112)

To return to the thought experiment that Smith introduced about a catastrophe befalling the inhabitants of some faraway country, the market connects even geographically separated communities. The cotton farmer in the United States would not be able to sleep easily tonight if a calamity befell the textile manufacture in China who purchased their crops. Thus, the market extends our circles of sympathy.

Building on Smith's theory of moral development, Paganelli (2010) has argued that the market is a civilizing space because it forces us to interact with a growing number of strangers. Because of this frequent exposure to strangers—who are socially distant from us and who cannot possibly fully sympathize with our passions—we develop the habit of moderating our passions so that they accord with spectators who know very little about us. The market, she argues, is a moral incubator because it allows for the level of social distance between individuals that is necessary for moral development. Storr (2018) builds on this insight: by enlarging our sphere of intimacy, the market places us into contact with an increasing number of potential friends—an increasing number of people whose mutual sympathy we genuinely crave and whose approbation we genuinely desire. Indeed, the relationships that we form within the market have a tendency to become overlaid with social content (Macneil 1986; Granovetter 1985; Storr 2008). “The market is, thus, a moral incubator,” Storr writes, “because of the volume and variety of individuals with whom we are brought into (potential) friendship” (2018: 467).

1.5 CONCLUSION

This chapter has argued that Adam Smith is an exemplar for how to incorporate ethics into our study of commercial society. In WN, Smith argues that markets are incentive-compatible with self-interest. In TMS, Smith argues that individuals not only pursue their self-interest, but also seek the understanding and approval of their peers. Together, these notions provide a fuller picture of the likely benefits and potential costs of our market dealings. Markets are certainly a source of wealth and are potentially a space of moral development (Muller 1993). But it is also important to recognize that markets work despite there being (1) limits to benevolence and (2) a potential for corruption.

This Smithian theory of markets that focuses on ethics stands in contrast to the narrower views of some champions and also some critics of markets. It is not uncommon for economists to describe the market as an amoral space and to analogize the market to

a tool (Langrill and Storr 2012). Like any tool, they often suggest, markets can be put to good and bad uses and can be used by good and bad actors. Just as a hammer can be used to build or to destroy a home, and a gun can be used to protect or to take life, markets can facilitate the distribution of life-saving drugs and of illicit drugs, can be used to procure uplifting books as well as pornographic materials, and so on. While agreeing that markets are tools, at least in some respects, Smith does not view markets as amoral spaces. Rather, he views markets as spaces where individuals live out their moral commitments, and that could lead to their moral corruption or moral development (Storr 2018).

Some critics of markets argue, however, that markets are not amoral spaces but actually promote and praise immoral behavior, such as greed, selfishness, and exploitation (MacIntyre 2007; Sandel 2012).¹⁹ If society becomes wealthier via the invisible hand of the market, they argue, it is at the expense of individuals being treated poorly by others or because they have become more immoral themselves. The material benefits of the market, goes this view, not only result in but depend on the moral decline of market actors. Smith at times seems to be endorsing this view. He acknowledges that markets depend on moral degradation, of at least some of us, and often result in collusion, exploitation, alienation, and the corruption of our morals. However, Smith's view of the morality of markets is not so one-sided. He also recognizes that the market is a moral teacher that rewards virtue and punishes vice. The student of ethics in a commercial society is neither a market critic nor a market defender but instead focuses on the ethical character of markets and market actors.

There is, fortunately, a growing literature that takes this Smithian approach to the study of markets that is mindful of the role of ethics in shaping human decision making, as well as the impact of market activity on our moral development. For instance, Gaus (2011) offers an examination of society, rules, and morality. McCloskey's trilogy of the bourgeois era (2006, 2010, 2016) examines the rise of the industrial revolution and argues that a shift in the view of commerce as a profession made it a respectable vocation and also encouraged the bourgeois virtues. In a series of works, Storr (2008, 2009, 2010, 2018) advances a theory of the market that not only distinguishes it as a setting for moral development but also as one that promotes and encourages honesty, trust, and friendship; Storr and Choi (2016) test and confirm this theory in the lab.²⁰

Furthermore, a Smithian approach to political economy not only highlights the potential dark sides of the market but also analyzes potential government reforms based on treating humans with dignity. Wolf (2017) examines Smith's critique of the British Poor Law and argues that it is in an ethical as well as economic analysis that has implications for modern-day aid efforts. Storr and Haeffele (2016) examine the potential for exploitation in the market as well as the limitations of government reforms. These works and others take up the challenge of a Smithian approach to political economy, though there is ample room for further research. We hope others take up this challenge as well.

¹⁹ For an examination of greed in Adam Smith's work and the scholarship on it, see Wight (2005). For a discussion on misunderstandings of Smith, see Rasmussen (2008).

²⁰ See also Badhwar (2008), Gardiner (1998), Henderson and Argyle (1995), and Hodson (1997).

REFERENCES

- Adler, Moshe. 1992. "On Being Honest and Behaving Honestly." *Games and Behavior* 4: 1–17.
- Akerlof, George A. 1970. "The Market for 'Lemons': Quality Uncertainty and the Market Mechanism." *Quarterly Journal of Economics* 84: 488–500.
- Ashraf, Nava, Colin F. Camerer, and George Loewenstein. 2005. "Adam Smith, Behavioral Economist." *Journal of Economic Perspectives* 19: 131–45.
- Badhwar, Neera K. 2008. "Friendship and commercial societies." *Politics, Philosophy & Economics* 7: 301–26.
- Berry, Christopher J. 1990. "Adam Smith: Commerce, Liberty and Modernity." In Peter Gilmour (ed.), *Philosophers of the Enlightenment Era* (Totowa: Barnes and Noble Books), pp. 113–32.
- Coker, Edward W. 1990. "Adam Smith's Concept of the Social System." *Journal of Business Ethics* 9: 139–42.
- Den Uyl, Douglas J. 1991. *The Virtue of Prudence*. New York: Peter Lang.
- Dwyer, John. 2005. "Ethics and Economics: Bridging Adam Smith's *Theory of Moral Sentiments* and *Wealth of Nations*." *Journal of British Studies* 44: 662–87.
- Fleischacker, Samuel. 2004. *On Adam Smith's Wealth of Nations: A Philosophical Companion*. Princeton, NJ: Princeton University Press.
- Fleischacker, Samuel. 2011. "True to Ourselves? Adam Smith on Self-Deceit." *Adam Smith Review* 6: 75–92.
- Force, Pierre. 2003. *Self-Interest Before Adam Smith: A Genealogy of Economic Science*. Cambridge: Cambridge University Press.
- Forman-Barzilai, Fonna. 2005. "Sympathy in Spaces(s): Adam Smith on Proximity." *Political Theory* 33: 189–217.
- Forman-Barzilai, Fonna. 2010. *Adam Smith and the Circles of Sympathy: Cosmopolitanism and Moral Theory*. Cambridge: Cambridge University Press.
- Gardiner, Coral. 1998. "Mentoring: Towards a Professional Friendship." *Mentoring & Tutoring: Partnership in Learning* 6: 77–84.
- Gaus, Gerald. 2011. *The Order of Public Reason: A Theory of Freedom and Morality in a Diverse and Bounded World*. Cambridge: Cambridge University Press.
- Gibbard, Allan. 1992. "Moral Concepts: Substance and Sentiment." *Philosophical Perspectives* 6: 199–221.
- Granovetter, Mark. 1985. "Economic Action and Social Structure: The Problem of Embeddedness." *American Journal of Sociology* 91: 481–510.
- Griswold, Charles L. 1999. *Adam Smith and Virtues of Enlightenment*. Cambridge: Cambridge University Press.
- Hanley, Ryan Patrick. 2008. "Commerce and Corruption: Rousseau's Diagnosis and Adam Smith's Cure." *European Journal of Political Theory* 7: 137–58.
- Hayek, F. A. 1978. *Law, Legislation and Liberty, Volume 2: The Mirage of Social Justice*. Chicago, IL: University of Chicago Press.
- Henderson, Monika, and Michael Argyle. 1995. "Social Support by Four Categories of Work Colleagues: Relationships Between Activities, Stress, and Satisfaction." *Journal of Occupational Behavior* 6: 229–39.
- Herzog, Lisa. 2016. "The Normative Stakes of Economic Growth; or, Why Adam Smith Does Not Rely on 'Trickle Down.'" *Journal of Politics* 78: 50–61.
- Hill, Lisa. 2006. "Adam Smith and the Theme of Corruption." *Review of Politics* 78: 50–61.
- Hodson, Randy. 1997. "Group Relations at Work. Solidarity, Conflict, and Relations with Management." *Work and Occupations* 24: 426–52.

- Hont, Istvan, and Michael Ignatieff. 1983. "Needs and Justice in the *Wealth of Nations*." In Istvan Hont and Michael Ignatieff (eds.), *Wealth and Virtue: The Shaping of Political Economy in the Scottish Enlightenment* (Cambridge: Cambridge University Press), pp. 1–44.
- Hope, V. M. 1989. *Virtue by Consensus: The Moral Philosophy of Hutcheson, Hume, and Adam Smith*. Oxford: Oxford University Press.
- Kumar, Nirmalya. 1996. "The Power of Trust in Manufacturer Retailer Relationships." *Harvard Business Review* 74(6): 92–106.
- Langrill, Ryan, and Virgil Henry Storr. 2012. "The Moral Meanings of Markets." *Journal of Markets & Morality* 15: 347–62.
- Levy, David M., and Sandra J. Peart. 2004. "Sympathy and Approbation in Hume and Smith: A Solution to the Other Rational Species Problem." *Economics and Philosophy* 20: 331–49.
- Loveman, Gary W. 1998. "Employee Satisfaction, Customer Loyalty, and Financial Performance: An Empirical Examination of the Service Profit Chain in Retail Banking." *Journal of Service Research* 1: 18–31.
- Macfie, A. L. 1959. "Adam Smith's Moral Sentiments as Foundation for His *Wealth of Nations*." *Oxford Economic Papers* 11: 209–28.
- MacIntyre, Alasdair. 2007. *After Virtue, Third Edition*. South Bend, IN: University of Notre Dame Press.
- Macneil, Ian R. 1986. "Exchange Revisited: Individual Utility and Social Solidarity." *Ethics* 96: 567–93.
- Marx, Karl. 1867. *Capital: A Critique of Political Economy, Volume 1: The Process of Capitalist Production*. Frederick Engels (ed.), Samuel Moore and Edward Aveling (trans.). Chicago, IL: Charles H. Kerr & Company (1915 edition).
- McCloskey, Deirdre N. 2006. *The Bourgeois Virtues: Ethics in the Age of Commerce*. Chicago, IL: University of Chicago Press.
- McCloskey, Deirdre N. 2008. "Adam Smith, the Last of the Former Virtue Ethicists." *History of Political Economy* 40: 43–71.
- McCloskey, Deirdre N. 2010. *Bourgeois Dignity: Why Economics Can't Explain the Modern World*. Chicago, IL: University of Chicago Press.
- McCloskey, Deirdre N. 2016. *Bourgeois Equality: How Ideas, Not Capital or Institutions, Enriched the World*. Chicago, IL: University of Chicago Press.
- Montes, Leonidas. 2003. "The Adam Smith Problem: Its Origins, the Stages of the Current Debate, and One Implication for our Understanding of Sympathy." *Journal of the History of Economic Thought* 25: 63–90.
- Montes, Leonidas, and Eric Schliesser (eds.). 2006. *New Voices on Adam Smith*. Abingdon, UK: Routledge.
- Morrow, Glenn R. 1923. "The Significance of the Doctrine of Sympathy in Hume and Adam Smith." *Philosophical Review* 32: 60–78.
- Muller, Jerry Z. 1993. *Adam Smith in His Time and Ours*. Princeton, NJ: Princeton University Press.
- Novak, Michael. 1984. *Freedom with Justice: Catholic Social Thought and Liberal Institutions*. New York: Harper and Row.
- Otteson, James R. 2002. *Adam Smith's Marketplace of Life*. Cambridge: Cambridge University Press.
- Pack, Spencer J. 1991. *Capitalism as a Moral System: Adam Smith's Critique of the Free Market Economy*. Aldershot, UK: Edward Elgar.
- Paganelli, Maria Pia. 2008. "The Adam Smith Problem in Reverse: Self-Interest in *The Wealth of Nations* and *The Theory of Moral Sentiments*." *History of Political Economy* 40: 365–82.

- Paganelli, Maria Pia. 2010. "The Moralizing Role of Distance in Adam Smith: *The Theory of Moral Sentiments* as Possible Praise of Commerce." *History of Political Economy* 42: 425–41.
- Pitts, Jennifer. 2017. "Irony in Adam Smith's Critical Global History." *Political Theory* 45: 141–63.
- Raphael, D. D. 1972–1973. "Hume and Adam Smith on Justice and Utility." *Proceedings of the Aristotelian Society, New Series* 73: 87–103.
- Rasmussen, Dennis C. 2006. "Does 'Bettering Our Condition' Really Make Us All Better Off? Adam Smith on Progress and Happiness." *American Political Science Review* 100: 309–18.
- Rasmussen, Dennis C. 2008. *The Problems and Promise of Commercial Society: Adam Smith's Response to Rousseau*. University Park, PA: Pennsylvania State University Press.
- Read, Leonard E. 1958. *I, Pencil: My Family Tree as Told to Leonard E. Read*. Atlanta, GA: Foundation for Economic Education (2010 edition).
- Roberts, Peter W., and Grahame R. Dowling. 2002. "Corporate Reputation and Sustained Superior Financial Performance." *Strategic Management Journal* 23: 1077–93.
- Robertson, Dennis H. 1956. "What Does the Economist Economize?" In *Economic Commentaries* (London: Staples Press), pp. 147–54.
- Rosenberg, Nathan. 1979. "Adam Smith and Laissez-Faire Revisited." In Gerald P. O'Driscoll Jr. (ed.), *Adam Smith and Modern Political Economy: Bicentennial Essays on The Wealth of Nations* (Ames, IA: Iowa State University Press), pp. 19–34.
- Rothschild, Emma. 2001. *Economic Sentiments: Adam Smith, Condorcet, and the Enlightenment*. Cambridge, MA: Harvard University Press.
- Sandel, Michael J. 2012. *What Money Can't Buy: The Moral Limits of Markets*. New York: Farrar, Straus and Giroux.
- Schliesser, Eric. 2006. "Adam Smith's Benevolent and Self-Interested Conception of Philosophy." In Leonidas Montes and Eric Schliesser (eds.), *New Voices on Adam Smith* (Abingdon, UK: Routledge), pp. 328–57.
- Sen, Amartya K. 1977. "Rational Fools: A Critique of the Behavioral Foundations of Economic Theory." *Philosophy & Public Affairs* 6: 317–44.
- Smith, Adam. 1759. *The Theory of Moral Sentiments*. D. D. Raphael and A. L. Macfie (eds.). Indianapolis, IN: Liberty Fund Inc. (1982 edition).
- Smith, Adam. 1776. *An Inquiry Into the Nature and Causes of the Wealth of Nations*, 2 vols. R. H. Campbell, A. S. Skinner, and W. B. Todd (eds.). Indianapolis, IN: Liberty Fund Inc. (1981 edition).
- Smith, Adam. 1982. *Lectures in Jurisprudence*. R. L. Meek, D. D. Raphael, and P. G. Stein (eds.). Indianapolis, IN: Liberty Fund Inc.
- Smith, Vernon L. 1998. "The Two Faces of Adam Smith." *Southern Economic Journal* 65 (1): 1–19.
- Storr, Virgil Henry. 2008. "The Market as a Social Space: On the Meaningful Extra Economic Conversations That Can Occur in Markets." *Review of Austrian Economics* 21: 135–50.
- Storr, Virgil Henry. 2009. "Why the Market? Markets as Social and Moral Spaces." *Journal of Markets & Morality* 12: 277–96.
- Storr, Virgil Henry. 2010. "The Social Construction of the Market." *Society* 47: 200–6.
- Storr, Virgil Henry. 2018. "The Impartial Spectator and the Moral Teachings of Markets." In David Schmidtz and Carmen E. Pavel (eds.), *The Oxford Handbook of Freedom* (Oxford: Oxford University Press), pp. 456–72.
- Storr, Virgil Henry, and Seung Choi. 2016. "Can trust, reciprocity and friendships survive contact with the market?" In Jennifer A. Baker and Mark D. White (eds.), *Economics and the Virtues: Building a New Moral Foundation* (Oxford: Oxford University Press), pp. 217–35.

- Storr, Virgil Henry, and Stefanie Haeffele. 2016. "Can the Government Stop Exploitation in the Labor Market?" Unpublished manuscript.
- Storr, Virgil Henry, and Arielle John. 2015. "The Impartial Spectator's Cultural Spectacles." In Charlotte C. S. Thomas (ed.), *Of Sympathy and Selfishness: The Moral and Political Philosophy of Adam Smith* (Macon, GA: Mercer University Press), pp. 32–52.
- Teichgraeber, Richard F. 1981. "Rethinking Das Adam Smith Problem." *Journal of British Studies* 20: 106–23.
- Thomas, Charlotte C. S. (ed.). 2015. *Of Sympathy and Selfishness: The Moral and Political Philosophy of Adam Smith*. Macon, GA: Mercer University Press.
- Viner, Jacob. 1928. "Adam Smith and Laissez Faire." In John Maurice Clark et al., *Adam Smith, 1776–1926* (Chicago, IL: University of Chicago Press), pp. 116–55.
- Werhane, Patricia. 1991. *Adam Smith and His Legacy for Modern Capitalism*. Oxford: Oxford University Press.
- West, E. G. 1975. "Adam Smith and Alienation: Wealth Increases, Men Decay?" In Andrew S. Skinner and Thomas Wilson (eds.), *Essays on Adam Smith* (Oxford: Clarendon Press), pp. 540–52.
- Wight, Jonathan B. 2005. "Adam Smith and Greed." *Journal of Private Enterprise* 21: 46–58.
- Winch, Donald. 1992. "Adam Smith: Scottish Moral Philosopher as Political Economist." *Historical Journal* 35: 91–113.
- Wolf, Brianne. 2017. "Beyond the Efficiency of the Market: Adam Smith on Sympathy and the Poor Law." In Peter J. Boettke, Christopher J. Coyne, and Virgil Henry Storr (eds.), *Interdisciplinary Studies of the Market Order: New Applications of Market Process Theory* (London: Rowman & Littlefield International), pp. 39–62.

CHAPTER 2

VIRTUE AND ECONOMICS, HORSE AND CART

JENNIFER A. BAKER

2.1 VIRTUE AND ECONOMICS

IN the early eighteenth century Bernard Mandeville thoroughly satirized the Aristotelian insistence that society would be better ordered if we only had more virtue. With his wry poetic portrayal of a beehive full of bees with human occupations, Mandeville asked us to just think for a minute: if “Honesty!” were to fill our “hearts” we could not possibly generate the flourishing economies that we do (1724: 211). Trinket sellers would discourage potential customers; grocers would sell at cost; tavern owners would see no demand. How many of us could remain employed? “Bare virtue can’t make nations live,” Mandeville writes. Not “in splendor,” but neither, it seems, otherwise (1724: 218). Once Jove puts honesty in the hearts of the bees, the “beehive” empties out, its economy caput and the bees left vulnerable to predation. Mandeville’s provocation is clear. We are not really designed for virtue, though it works fine as a cover for what we are really good for: commerce.

There is a lot at play here: what virtue is, what we think is good in lieu of virtue, what is required of us to bring about a functioning economy, what a functioning economy even is. Mandeville’s misinterpretations (getting virtue wrong by pitting it against self-awareness and concern for consequences) exaggerate tensions, but he does not create them. Aristotle’s own work is evidence enough that the relationship with economics and ethics has been awkward from the start. In this chapter, in an attempt to defend the relevance of virtue to economics, I want to answer three questions that, if left unaddressed, would seem to stymie any effort to think about ethics and economics at once.

First, what is a moral agent going to need to know about economics? Economics is the study of what is not immediately intuitive for anyone, the “subtle, difficult to detect, and

hard to understand” routes to desirable social outcomes (Dasgupta 2003: 70).¹ Most economists think that we must approach these phenomena with mathematical modeling to begin to understand them. Surely moral agents cannot be required to be economists themselves.

Second, how and why should an agent’s own moral reasoning reflect economic analysis, if it can even be put in a form that is communicable? Non-economists do not always realize, and may simply not see, that we operate under various constraints with varying access to resources, skills, and information. We make choices without being aware of information we would count as relevant or even how our choice sets themselves involve factors of which we remain unaware. And even if we could be supplied information about incentives, opportunity costs, and selection effects relevant to our case, this is not the same as information about the actual “routes” to whatever outcomes, as these change with the choices others are making.

Finally, how should we determine when ethical considerations ought to take priority over economic rationales? Any type of philosophically detailed ethical approach might provide workable answers to these, but I want to show off what virtue ethics can do to provide an answer. And I do not mean what virtue ethics can do prospectively, but what we can do with it given elements of the view that already exist.

It might seem silly, or arrogant, for someone from a small field like virtue ethics to suggest we can wrap an explanation around a tremendously large, innovative, and productive field like economics. But virtue ethics is good at some things that economics does not even attempt. We need virtue ethics in order to understand *our own* moral agency, and it is through our moral agency that we process and understand the discoveries of economics. Our outlook as agents is ever-present, no matter how much assistance it might require in understanding the unduly complex and uncharted larger world. Virtue ethics addresses agents in their own language concerning what is good and what is worth pursuing, and also introduces concepts crucial to representing the values that economists study, several of which I point to in this chapter.

The benefits of maintaining clear distinctions between the fields go both ways. Virtue ethics can play a crucial role alongside the robust practice of economics, as it maintains a standard for the justification of moral claims; and economists describe the world in which we must act, as best we can. It is therefore important that we not confuse imagining a society run by virtue with an argument against virtue itself.

Friedrich Hayek has attempted to defend Mandeville from this mistake, by arguing that Mandeville’s focus on “vice” is “but a special case of a much more general principle” (1978: 253). The emphasis that “had provoked all the moral indignation” is “almost

¹ Gary Becker has an oft-cited description of economics, but it is not aimed to help agents themselves understand the point of economic exchange: “The combined assumptions of maximizing behavior, market equilibrium, and stable preferences, used relentlessly and unflinchingly, form the heart of the economic approach as I see it” (1976: 5). Earlier, Lionel Robbins described economics as “a science that studies human behavior as a relationship between limited resources and unlimited wants which have alternative uses” (1932: 15).

irrelevant” to Mandeville’s lasting methodological insights, Hayek argues. Mandeville’s “main contention” was that

in the complex order of society the results of men’s actions were very different from what they had intended, and that the individuals, in pursuing their own ends, *whether selfish or altruistic*, produced useful results for others which they did not anticipate or perhaps even know. (Ibid., emphasis mine)

With this slight update of Mandeville, one can see why he is sometimes given a nod for originating a very early version of today’s economic methodology. For us to leave ancient assumptions about social order behind, economics needed to operate at the level where the intentions and motivations of agents, *whether selfish or altruistic*, are not of direct relevance.

There is an “ethical level” of analysis, a matter of making distinctions between motivations, differentiating selfishness from altruism on some type of reasoned (and, as is required today) empirically-sound basis (Russell 2009).² There is also an “economic level” of analysis, concerned with how beehives work, and only derivatively the bees. Such an analysis might still tell us a substantial amount about bees. For example, behavior was once subsumed under the heading “self-interested” by economists. By their own lights some now find “generation of trust” or “indirect reciprocity” to be more descriptive (Fehr et al. 2002; Gintis 2000; Henrich et al. 2005). Even economists wanting to understand choice “from the inside” are working at the economic level, since their interest is our behavior in the aggregate, a matter of statistical regularity (Rhonheimer 2015). Don Ross argues that economics ought to only be done at this level because it is the pursuit of “*general* knowledge about markets” (2014: 5, emphasis mine). He explains that economic analyses do not become more “powerful” if done at the ethical level, because issues of agency (such as how agents frame a problem) lack “even a relatively general solution either theoretical or instantiated” (ibid.). Rather than attempting to figure out our moral psychology, an economic analysis is going to view ethics as a kind of commons problem (Frey 1997).

The mere suggestion that there can be a level of analysis where it does not matter what agents intend, where some sort of invisible hand or emergent order can work with participating agents’ motives whatever they might be (so that beneficial outcomes are “motive independent”), can seem anathema to common understandings of ethics. And economists can find types of “value neutral” positioning in their field suspect (Sen 1987; Hausman and McPherson 1996; Phelps 2013). The accusations include that economics is already attached to many ethical assumptions and that these are merely shielded from scrutiny if we pretend work exists above the need for ethical debate over the common good (Anderson 1995; Alvey 2011; Schliesser 2016). There have been recent worries that

² Aristotle’s economics, which opposes retail trade and assumes commodities have an objective value we can discern if we are philosophical enough, has been disconnected from his ethics since even ancient times (Meikle 1995).

taking an economic approach to “ethical behavior” may make the mistake of looking to explain a phenomenon before coming to understand it (Baurmann and Brennan 2016). As an example, Michael Sandel argues that applying the notion of scarcity to ethics “ignores the possibility that our capacity for love and benevolence is not depleted with use but enlarged with practice” (2012a).³ He has also pressed the case that the damage done extends outside of economics, so that its pretenses damage our appreciation of moral reasoning itself (2012b).

Virtue ethics is a possible solution to each of these worries. Its account of practical rationality can help us conceptualize the ethical and economic “levels” of analysis and allows us to move back and forth between them, while keeping mindful of the difference between our own moral agency and the purpose of economic analysis. This suggestion is not similar to the urging of economists to “infuse” economics with general ethical concerns (a common, though painfully vague, recommendation). Nor am I asking for anything like a return to the political economy of the eighteenth or nineteenth century (Schliesser 2016). To give up all that economics has discovered since seems as foolhardy as seeking a return to Aristotle or Augustine’s economics. Suggestions that we incorporate general “ethical concerns” into economics can amount to a rejection of specific and fully fleshed accounts like virtue ethics (although I argue how this is not necessarily so below). Any ethical theorist will, of course, insist that the ethical theory we choose to use actually matters. (In this case, it will be the account of practical rationality, shared between various types of virtue ethics, on which my proposal depends.)

Neither do I envision a kind of “business ethics” for economics, where ethicists are asked to weigh in on the field’s practices by debating ethically relevant conceptual issues. Such efforts can be flattering to ethicists, as they are granted “special expertise in the ‘logic of concepts’ and thus a unique responsibility for determining universal norms” (Ross 2014: 5). And yet ethicists debating each other over concepts they have invented—for example, there is a debate in the field of business ethics over what business ethicists mean by profit—seems impractical. It is far better, I think, to recognize that economists, in “the course of empirical discovery and theoretical refinement,” are capable of making their “technical concepts more coherent and consistent” (*ibid.*). And surely enough, methodological debates within the field of economics are lively and productive, currently including how behavioral economics is to fit into neoclassical economics, whether methodology should be individualist or not, how psychologically grounded accounts of preference ought to be, and the degree to which ethical motivations could factor into modeling (Angner and Loewenstein 2012; Backhouse and Cherrier 2017; Chetty 2015). I am not interested in virtue ethics being interposed into any of this work. Commitments, trust, non-selfish motives, reciprocity, intrinsic motivation, inequality aversion, and fairness: these and more have now been identified through economic methods. To have

³ Economic approaches are structured by the assumption of scarcity. Virtue is not scarce. It is, notoriously, uncompetitive. Aristotle has us imagining two virtuous friends as rivals over only who would sacrifice for the other (Annas 1993: 252). Even this is bound to be misunderstood by those committed to the idea that we are only self-interested, even when we do good.

attached ethical assumptions to this research before it was done could have cost us these findings. And the independence of these findings is of value to philosophy generally. We do not want to put the cart before the horse. Of course, I realize, after mentioning all of this work that economists are doing on ethical behavior, I need to demonstrate why we still need ethical theory (the horse).

2.2 WHAT IS VIRTUE?

“Virtue” means far too many things when unattached to an ethical theory. Deirdre McCloskey (2006) has provided a rich description of the market-friendly “bourgeois virtues”—pace Mandeville, she argues that the prime market virtue is indeed honesty—but she only intends these to refer to popular mores. She makes it clear that she is working at the economic level, and writes, “virtue in each person is strictly speaking neither necessary nor sufficient for the virtue of and certainly not the survival of the whole community” (ibid.: 247). Virtue “in each person” is what the theory of virtue ethics aims to assess. There is a difference in acting like an upright bourgeois and having “virtue itself.”

What is virtue itself? Julia Annas has perfected a quick but accurate answer:

Virtue is just the virtues, admirable traits of character like bravery and justice, united by the fact that they share good practical reasoning about what should be done. A virtue is a disposition, that is, a habit of acting which has been built up through practice, though it is never thought of as a mindless habit, since it is a disposition to deliberate and to make decisions. (2001: xix)⁴

The “point of virtue,” she explains, is that we “learn to think” for ourselves “about ethical matters” (about which everyday thinking is usually quite wrong!) (ibid.). Virtue ethics is what we term the type of ethical theory that descends from the ancient Aristotelian account of our moral psychology (a normative account) and provides criteria for determining particular cases of right and wrong (Annas 2011). Virtue ethics admits of many diverse approaches. Some virtue ethicists work at a conceptual level, with a focus on particular virtues. The type of virtue ethics I will refer to is what I consider “traditional” because it takes practical rationality to be necessary to virtue. There are even many varieties of virtue ethics among these.

Sometimes economists who are interested in philosophical accounts of ethics work at a bit of a remove from the details of any particular account. We see this happen with the

⁴ Charles Griswold describes Adam Smith maintaining the same thing the ancients did when he explains, “moral rules are not . . . being *applied* at all; rather, in light of general maxims, an evaluation of the specific context is undertaken and a decision reached” (1999: 190). These maxims (or precepts or norms) have a “protreptic and pedagogic role,” and there is “no single procedure, method, or formula of moral justification or reason giving, even though we can specify the general structure of reason” (ibid.).

work of Adam Smith. Smith might get interpreted as promoting self-interest (Stigler 1975: 237) rather than reason and understanding and self-command (Sen 1987: 22). Economists can be found puzzling over why Smith would ever think beneficence would be performed without interest in reciprocation (Hanley 2015: 106). Smith scholars argue that his is a traditional account of virtue which recognizes that, when virtuous, our practical reasoning has developed to such an extent that our interest in virtue is no longer instrumental (Griswold 1999: 190; see also Smith 2009). Virtue cannot be easily assimilated to “self-interest” because virtue becomes the way developed agents self-identify, which changes the way they value everything else (Becker 2017; Annas 2011).

Working at a remove from the details of particular accounts can also result in “virtue” being assimilated to rather religious notions of good behavior. Those invoking virtue in this way might refer to a “pure virtue ethic” and suggest that, for someone with “pure” virtue, the actual consequences of her actions are not her concern (Baumann and Brennan 2016). It can certainly *sound* as if an agent acting “for the sake of virtue” has nothing else in the balance, but the painstaking process of becoming virtuous involves paying attention to the impact of our choices. The practical rationality necessary to virtue is a matter of coming to understand what the right goals of our behavior ought to be, and then coming up with a plan to ably pursue those (see Aristotle, *Nicomachean Ethics*, VI. 5, 1140a25–31, b4–11). By definition, if it causes harm, it is not virtue. Socrates explains this as early as Plato’s *Euthydemus* (around 380 BCE).

We have just corrected Mandeville’s take on virtue in two ways: it is not just disguised self-interest, and it cannot be virtue if it leads to social disaster. A third suggestion we can find in his poem is that virtue is largely a matter of self-delusion. And of course research in social psychology surely underpins concerns about over-inflated assessments of our personal ethics. But virtue ethics based on practical rationality is a serious critic of the typical way in which we regard ourselves. Given research on how easily we self-deceive, it may seem that any ethical approach associated with “know thyself” has grown outdated. But many of the insights social psychologists are now tracking carefully are in line with why virtue was urged in the first place (Russell 2009). This passage from Epictetus is a good correction to the idea that “know yourself” means something scientifically implausible: “if someone orders a singer in a chorus to know himself, would he not attend to that order by both paying attention to his fellow chorus members and to harmonizing with them?” (quoted in Johnson 2014: 12). In other words, what we must “know” according to traditional virtue ethics is not something ineffable; virtue is built out of practical rationality, which is a matter of sorting through various candidate moral norms.

Contemporary virtue ethicist Lawrence Becker (2017) emphasizes that practical rationality is a matter of testing, rejecting, and endorsing normative propositions about our projects, considered all together. As he explains, healthy agents will acquire “strong norms” corresponding to the usual notions of wisdom, justice, benevolence, beneficence, courage, temperance, and other traits that are “standardly called virtues” (124). We are to look for clear instruction and self-designed rules for our behavior, ones that can be promulgated between agents and help us to become aware of the values we can associate with our choices.

One of the foundational assumptions of traditional virtue ethics is that failing to examine, understand, and endorse the basis of our behavior is going to lead to misunderstandings that affect our general ability to be practically rational. This, in turn, negatively affects the efficacy of our agency, altering even our ability to issue sound judgments or policy. This is why virtue ethics recommends that we sort through our commitments and use practical reasoning until we find an actual norm that might integrate the good of promoting general affluence into our general understanding of what is good. Elizabeth Anderson (2015) has pointed out that typical and inadequate “takes” on economics are used as (poor) justification for unjust policies. This is exactly what a traditional virtue ethics would target: the things we think are true about economics and the way we use these. In the next section, I explain how economics, which helps us track outcomes, has relevance to our own personal moral reasoning. But we should never regard our personal moral reasoning as some proxy for economic thought.

2.3 THE ECONOMIC WAY OF THINKING

When it comes to our economic behavior, traditional virtue ethics would have us look for what the ancients called a “sayable,” some way of putting into words the point of economic production and exchange (Fortenbaugh 2002). Previously, I proposed a norm or precept that I thought might work to guide agents through what can be ethically puzzling about markets (Baker 2009). It was tentative, and I put it forward after rejecting more common contenders, such as “pursue self-interest,” “conserve resources,” and “cooperate.” An economist following along had a friendly suggestion I had failed to consider, which was briefer, more direct, and, to him, obvious: maximize surplus.

This is certainly *a way* to relate the two levels we discussed: just compress them into one. Economics would provide the content for even very personal guidance. If economic explanations replaced ethical ones, the idea goes, we might generate far more surplus. An economic approach offers an alternative to the stories we tell ourselves about agents and their good behavior (this is unlike Mandeville’s description of virtue as a useful ruse). One might think that taking an economic approach to ethical behavior is to examine it as it really is, avoiding the pretense that we are better than we are or motivated more nobly than we are (Simler and Hanson 2018). I want to argue against these ideas. You may, of course, already agree that “maximize surplus” is a poor guide to personal ethics. But if I take care to make my case, it should work as both a defense of virtue ethics and a way of clearing the ground for an approach, based on practical rationality, to the ethical relevance of economics. The periodic efforts of economic columnists to rehabilitate Charles Dickens’s Scrooge are an example of what is lost if we fail to make the distinctions I am after.

Every December, you can find an economic columnist jauntily suggesting that Charles Dickens’s Scrooge has been miscast as the season’s villain and woefully misjudged. Perhaps these columnists are best thought of as exaggerating for effect. They are usually

advocates for the economic way of thinking or, if economists, writing in a for-the-public mode (Waldfoegel 2009; Cowen 2013; Landsburg 2004; Harford 2015). But they give us an opportunity to use Scrooge as a test of the idea that “maximize surplus” could work as a guide to personal ethics, and to explain why a “no” answer helps demonstrate virtue ethics itself.

His defenders take pre-reform Scrooge, the brisk businessman who is trusted to keep accurate accounts, to be following the market’s logic, which is the path to general welfare. Scrooge is better able to follow this logic precisely because he is unattached to traditional moral categories on which so many of us unproductively focus. The columnists explain that, no matter how much sympathy we want to muster for his circumstances, Bob Cratchit is psychologically no different than the rest of us, and is merely maximizing his opportunities, just like Scrooge. After all, had Cratchit been able to be employed for more elsewhere, he would have left Scrooge’s business. Scrooge is not only blameless for the Cratchit’s family’s struggles, he had provided for them as much as was possible. Pre-reformed Scrooge was enabling the welfare of others, seeing the right “routes” to good social incomes, the ones most of us cannot appreciate (Landsburg 2004).

What about Scrooge’s personal attitudes? Dickens characterizes Scrooge by calling him “a squeezing, wrenching, grasping, scraping, clutching, covetous old sinner” (1843: Stave One). These are either forgiven or lauded in light of Scrooge’s productive capacities—or his defenders merely think such attitudes are minor matters, impossible to tally given the importance of economic outcomes. Some of his defenders go so far as to take up Mandeville’s ethical reversal and argue that Scrooge was a better man before his transformation, due to his previous economic contributions outweighing any turkey-giving (Landsburg 2004). Miserliness is actually generosity! Overturning common sense still reads strangely, as you can see from being told that Scrooge’s “not spending” benefits the rest of us more than his spending:

What could be more generous than keeping your lamps unlit and your plate unfilled, leaving more fuel for others to burn and more food for others to eat? Who is a more benevolent neighbor than the man who employs no servants, freeing them to wait on someone else? (Ibid.)

It is as if the Scrooge-defenders, recognizing that economics deals with the “subtle, difficult to detect, and hard to understand,” have come to expect that people will be mistaken on all sorts of other topics as well. They are missing that a compromise can be struck between those pushing a purely economic story about Scrooge and those advocating for an account of virtue, if we only recognize distinct ethical and economic levels of analysis. Then, no eliminativist approach to ethics will seem necessary to the economic point that I believe these writers are actually attempting to make. As “not spending” does not even approximate “maximize surplus” in terms of being a description of sound economic behavior, it is charitable to assume that defenders of Scrooge are attempting to get us to appreciate some points about economics rather than ethics. I think those who defend Scrooge are defensive about the good done by those who transact in the market, whether

spending, saving, or loaning. It seems that these authors are worried that traditional ethics will give moral credit for failed efforts to contribute to the general good, as if virtue will “crowd out” sound economic behavior. (Sandel’s worry, of course, is the other way around.) Perennial arguments that these economic columnists make against gift-giving are similar.⁵

But pre-reformed Scrooge ought to be no sort of role model, as Dickens, in his memorable characterization, gives us three “red flags” signaling that pre-reform Scrooge, for all his cleverness, is not practically rational and is suffering for it. First, if Scrooge were practically rational, he would not feel afraid to confront his memories or indeed others’ challenges. Such confrontations expose and generate the very content on which practical rationality is to do its work. We are to aim at interpreting no memory as being at odds with what you regard yourself as committed to; if you have a successful justification of your behaviors, no query from another person is supposed to make you self-conscious about your approach. Second, pre-reformed Scrooge’s rage is a sign of a failure of practical reasoning. Finally, his empty rhetorical responses to being challenged—such as being asked, “Are there not poor houses?”—demonstrates a failure to be able to defend commercial behavior in common parlance, recognizing it as a way to contribute to the common good.

Let’s consider how better practical reasoning could change Scrooge, rendering him able to self-examine, less full of rage, and capable of giving a plausible account if asked to justify his behavior. Again, unlike those who merely do conceptual analyses on the virtues, a traditional virtue ethicist will not just recommend he become kind or generous, with the assumption that we all know what that means. Instead, the focus is on developing one’s practical rationality (Nussbaum 1988). The emphasis is on allowing agents to test the norms that they adopt, both those they pick up from others and their culture, as well as those they generate when they follow norms they cannot justify or that poorly match their behavior. Scrooge, in other words, ran his own kind of experiment, attempting to live by “sayables” that could not capture value he recognized in his life otherwise. He was testing out the “selfishness axiom,” that is, “the assumption that individuals seek to maximize their own material gains [...] and expect others to do the same” (Henrich et al. 2005: 797). His sputtering confusion is, again, the exact kind of consequence an account of virtue based on practical rationality anticipates when one has not

⁵ Economists writing on this topic (for the public) explain that what is lost in the transaction is not individual profit but an opportunity to maximize surplus. These critics of our cherished habits explain that we are wasting 18–20 billion dollars a year by buying gifts that recipients would not buy themselves. Economists of course mean something different by “efficient” than the general public does, but writing for the public, you can find these authors conflating the two senses. Gift-giving is presented as a “negative trade” likely to “destroy value.” They put it this briefly: due to “information and motivational effects” we know we ought to be give money so others can buy their own presents (Cowen 2013). Surely enough, the critics of gift-giving tend to segue into what might be their actual aim: a critique of inefficient charities. This is the concern they can justify. If an agent gives a gift to a friend, we need not do anything like “set economics aside.” Instead, I am arguing that we should evaluate her behavior according to standards like those inspired by virtue ethics.

considered whether their commitments and behaviors are justified. We are to revise these until we can ensure these justifications seem coherent to ourselves and others.

Because we are to personally endorse these standards, they should not be rendered artificially technical, as if they were generated remotely. They really should be the things we begin to say and reflect the kind of self-talk in which we engage. Accordingly, the articulated standards might look like descriptive claims rather than explicitly ethical recommendations. As Seneca writes in his *Moral Letters to Lucilius*:

Some consideration may well be in our minds, but yet not sufficiently accessible; we begin to get at them only when they are spoken aloud. And other things are there, but scattered in widely separated (mental) locations, and the untrained mind cannot bring them together. So we have to bring together and juxtapose such considerations, so that they can have more impact and do more good to our character.

(Letter 94, quoted in Inwood 2005: 118)

Articulating our current assumptions gives us the material with which we are to work, work that begins with that which we can psychologically access. Articulation of standards becomes a prompt to think through a claim's appropriateness and what its adoption into an explicitly endorsed ethical norm might harm or improve. If we articulate a norm that better matches our intentions than it does our actual behavior, we can then become motivated to change our behavior. Or we might instead approve of our behavior and reject a proposed norm.

Engaging in this iterative process is how virtue ethicists expect us to improve our understanding of our own agency. And those who continue to act in unjustified manners become a source of information to us as well, as their inability to offer a justification is evidence that can be expected to coincide with unhappy psychological effects. The standard of intelligibility is easily achieved by agents themselves in terms of stated norms that accurately describe their commitments. The ancient examples show us that the approach we can take to this understanding is not loftily abstract (Baker 2013). In fact, abstract meditations on particular virtues are unlike the examples of applied ethics from the ancient world. Then, the practical focus was often a matter of addressing the public in an attempt to identify the right norms or standards to follow. Finding the right norms will have a "favorable impact" on our characters (Inwood 2005: 119). Failing to find them is expected to do the opposite.

The ancient virtue ethicists put into practice a simple approach: a matter of questioning. It is a method like that associated with Socrates, and would also be part of what Seneca describes as the "procedures of choice" (ibid.: 112). First, an agent must be encouraged to think through what counts as good or ethical. Again, to articulate a norm is one way to acknowledge one's thoughts on an issue. Second, an agent must distinguish between the claims she reflectively endorses and those she follows in her actual behavior. Third, she must test the fit between claims she articulates in a manner that seems rather standard: Are these coherent? Are they factually accurate? Do they reflect prejudice or bias? How would you react if you were treated in accordance with any norms you are

considering? Finally, your claims must be reconciled with other claims or norms we consider endorsable, as well as with other motivations we have (Baker 2013; Lutz 1999).

This type of questioning can reveal problems with the beliefs that guide our behavior. For example, if we engaged in this approach when giving to charity, we could end up aware of any lack of justification for our behavior, just as Scrooge would have when refusing to give to any charity. We would be checking the rationale against commitments that are already central to our self-understanding. Can we be miserly and virtuous? No. Why not? Attaching unexamined importance to one's own earnings is at odds with practical rationality. The possibility, of course, of being a miser is an ancient one, and diagnosed as a tempting option by Aristotle himself: money, in particular, as it represents good, can come to have an attraction as if it were itself a non-instrumental, unlimited good. The kind of addictive quality to money accumulation is due to this, he argues in the *Politics* (Meikle 1995). Any type of goal without limits is always mistaken, Aristotle explains, as it means it is not being subject to a defensible rationale. The psychological proof is that you will come to pursuit of such a goal as regrettable precisely because it was (to you) kept indefensible, as in Scrooge's story. Scrooge, without a ready defense of his focus on just business interests, alienates himself from both the things he could defend as good and from his own practical rationality.

A feature on CNN on the collected "regrets" of former debt collectors illustrates this idea. As one interviewee explains,

If someone told me they only had \$150 to their name this month and needed to feed their entire family of five, I would say, "I don't care, this is your fault and you owe the money. Pay it." I would even use blind threats like "I know where you work" to intimidate them. It was survival of the fittest, and I thought being aggressive was the only way to succeed. And then it became an addiction—being authoritative and abrasive was like a high. (Harsh 2010)

This short report really captures the description of unjustified behavior offered by traditional virtue ethics. Virtue ethicists would focus on these elements of this man's description: the norm of "survival of the fittest" being associated with market success; market success becoming something more than a means to an end (the addictive feeling); and the tension (now revealed by the interviewee) between the man's normal concern for others and his on-the-job disregard for their dire situations. The rationale of "you owe the money" was not working for him, and he "can't wait to get out of this line of work" (Harsh 2010). And as I hope is becoming clear, virtue ethics would suggest that the man rid himself of his current beliefs that are in conflict with other things he endorses and making him feel ashamed of his work. I would suggest that his work can be defended if he only alighted on the right account of it (and that, then, would be encouragement for him to modify his vicious approach). But we do not have, on hand, accounts of personal economic norms. Of course, we have endless feedback from those with whom we are working about how to behave, but these norms are not enough to satisfy the person questioning why he is doing what he is. Scrooge is in the same situation when visited by his partner's ghost.

What might help? An articulable, commonsense explanation of the good of the market. This should be a norm (or maxim or precept) that can be used in our own deliberation. As an example, I have suggested the following:

We are behaving justifiably in the market when we are aware that it is the right time and place to treat resources as scarce for the sake of contributing to a market system that generates information that can be used by others immediately, and long term as a means to general affluence. (Baker 2015)

This candidate norm can be understood without reference to concepts that stand in for unclear philosophical accounts; it is descriptive of what actually happens in a market transaction; and it is capable of being guiding when it comes to the actual choices in the market in a way that a commitment to values such as freedom or efficiency are not. A norm like this matches our behavior in a way that some offers do not (we know we are not really feeling benevolent when we try to upsell a client). And it does not require that we maintain the selfishness axiom. Note, however, that in asking us to generate standards for our own agency and the justification of moral claims, virtue ethics leaves room for us to act out of self-interest (and expect the same). The norm also acknowledges that, when in a market mode, following market norms, we treat commodities as limited resources. We bid and barter. We do not give them away. A reminder that such market norms are chosen, justified but capable of being rejected on a case-by-case basis, is a reminder of the choices we all have, even when in markets. For example, our role in a large market is dispensable, given sufficient affluence around us. This alone has the potential to change attitudes that interfere with our self-conceptions. Not every sale Scrooge makes matters to general affluence! The urgency is all his.

Finally, we need a norm we can follow with consistency given our other commitments to what is good. The loan collector feels guilty for what he has done; his methods are clearly not in line with his beliefs about what people deserve. Yet he also seems to think his efforts are wholly unjustified. I say this as a person who has received calls from debt collectors: it is not an occupation without a role in promoting general affluence. A norm that explains the connection between our focus on conserving resources can help to justify occupations that have been casually treated as beyond the pale ethically.

The potential for general affluence is the good you are promoting when you make a deal in the market, and we need not think of ourselves as only doing this indirectly or underneath some type of delusion. You contribute directly by committing to the norm of conserving the resources under your current or future control. The underlying concern of some of the defenders of a pre-reform Scrooge are a way to highlight the social good of general affluence, an effort to criticize well-meaning efforts that, in some way, backfire given the original expectations. Yet giving inefficiently to charitable causes is something traditional virtue ethics has us critique. It is two things: a failure of self-examination, as one cannot justify the behavior, and also a failure of practical rationality, as it ignores the role that economics ought to play in what Amartya Sen has called the “engineering” aspect of pursuing social outcomes (1987: 4–7). Once we have agreed upon an end, the

proper means are of relevance, certainly when the project is a social, public one. Making use of virtue ethics obviates the need to defend Scrooge as a good guy in order to make this ethical critique.

Unlike notions of moral “purity,” traditional virtue ethics has been able to defend the pursuit of profit for some time. Of course, Aristotle thought the pursuit of profit was incompatible with virtue, but this was due to his own economics which insisted that profit only resulted from confusion about value. Virtue ethicists that came right after him (such as the Stoics) argued for the compatibility, likening the pursuit of profit to any other prudential skill (Natali 1995, 2003). Contemporary virtue ethicist Christine Swanton has put forward a virtue of productivity, arguing that we ascribe too many aims to business activity and ought to think of it as having a focused and justifiable goal of profit-making (Swanton 2016; see also Munger and Russell 2017). And economists themselves have begun listing very plausible potential virtues that suit market behavior. Luigino Bruni and Robert Sugden (2013) offer us enterprising alertness, respect for other’s taste, acceptance of competition, self-help, stoicism over reward, trust and trustworthiness, and a kind of universalizing towards other agents. The limitation on pursuing profit, for a virtue ethic, will be that it be done with a defensible rationale. That’s all. But, if this defense were put in the context of overall human and social good, this requirement, put to each of us, could be as powerful and reformatory as any idea at all.

2.4 FITTING VIRTUE TO GENERALLY ETHICAL BEHAVIOR

I have argued that the use of ethical theory can help economists to make points about socially beneficial behavior, but, as I have mentioned, economists are also doing work on topics, such as motivation, that should surely be counted as “ethical.” Are philosophical accounts of ethics and new discoveries about our ethical motivations competitors? Virtue ethics has long had a resource for explaining why they are not. The notion of a general, “normal” amount of ethical development has been considered something other than “virtue” since ancient times. Often thought Stoic, it is now clear from Arius Didymus’s discussion that it is an Aristotelian notion as well (Görgemanns 1983). To use the ancient terminology, economists today assist us in explaining standard moral behavior or what ancient virtue ethicists called *oikeiosis*. This is a revisable, science-based account of standard psychological development. It is what ancient virtue ethicists expected to account for our becoming social and loving people, something that happens to most of us whether we have ever heard of virtue at all.⁶ An account of virtue fits itself to this data. Today we can think of *oikeiosis* as what we learn about behavior working at

⁶ See Becker (2017) for a list of empirical, psychological preliminaries to a virtue ethic.

the economic level. In other words, using virtue ethics alongside economics in no way requires economists to aspire to treat ethical subjects in as value-neutral a way as some behavioral scientists do (for example, those who study motivation, such as Ainslie 2013, seeing addicts as examples but not negative ones).

There is yet another concept from virtue ethics that leaves room for economists to do what comes naturally when discovering behavioral patterns that lead to good social outcomes: recommend behavioral changes. Without expecting economists to learn or care about virtue ethics, we can still suggest that their recommendations for us be categorized, by a virtue ethic, as an “appropriate action.” (Again, the Stoics are known for this distinction, but Arius Didymus records the term being used by early Peripatetic or Aristotelian philosophers; see Fortenbaugh 2002.) The recommendation of an “appropriate action” is fodder for a virtuous agent’s own practical reasoning, which determines whether to follow it or not, but is also guiding and usually important to reference in one’s own reasoning. Appropriate actions might (and might not) be what a virtuous person would do; what matters is that they can be recommended apart from considerations of agential moral psychology. They need not be done as the virtuous person would do them. If we maintain a distinction between virtuous actions (which require not just doing good but also the right reasons and motivation) and “appropriate actions,” we will keep from putting undue expectations on economists studying and recommending behavior. They can do this without recommending or interfering with virtue. The category of “appropriate actions” gives economists (and others) the chance to be explicit about what grounds their recommendations. There would be no pressure to defend these as ultimately or exclusively ethical.

And so, we should acknowledge that an economist can show that a type of behavior encourages trust, and our reaction may still be, “why should I care?” Virtue ethics comes to the rescue here, doing as much as it can to argue for good behavior. But it is not just a handmaiden to economic recommendations, of course. Virtue ethics also stands as a reminder that we have options and that no suggestion, no matter how prudential, replaces the standards we devise through our own commitments to ethics.

There is one final concept from traditional accounts of virtue that I think can help us to maintain the distinctions I have argued for above, as well as underscore how and why it is that a practically rational agent might reject or endorse general market norms in a given situation. This is not part of any Aristotelian account, and is distinctively Stoic: the notion of “moral indifferents.” For Aristotle, some goods external to virtue were considered necessary to a happy life. The Stoics worked out a two-scale value schema so that we need not think of virtue (the development of practical rationality) as somehow commensurable to the regular goods we pursued and needed in our lives. This distinction is, itself, a bit tricky, but only initially.

The term “indifferent,” for example, was meant to be provocative, to shake us out of our casual assumptions, because indifferents are, of course, matters of life and death, well-being, and scarcity. We regard these as so valuable that we easily choose them over virtue, but we should not. As Sharpe explains,

problems emerge when we prioritize achieving some external outcome over practicing virtue, as if this external were a good necessary to our happiness, so that we become willing to compromise virtue in order to achieve this external outcome. (2014: 35)

We need vivid reminders that these goods are not the same as virtue, and so “indifferent” is proposed as a second category of value.

This concept can assist agents aiming to develop their practical rationality, leading them to regard concern for their agency (their developing practical rationality) as a limit on both the pursuit and the valuation of any particular “good” by categorizing it, in relation to one’s agency, as a mere “indifferent” (Long 2002: 110). A.A. Long explains that for the Stoics,

when someone says something is good for them, we should perhaps rephrase that by saying that it has value for them, but is not strictly speaking “good.” Food, water, and shelter all have value to me, but they are not “good,” for only virtue is good.
(Ibid.: 24)

The Stoics were likely inspired by the ancient Cynics, who denied anything but virtue could be counted as good at all. But one of Stoicism’s founders, Plutarch, writes, “They are insane who make wealth, health, freedom from pain, soundness of body of no account and who do not cling to such things” (quoted in Long 2006: 347). As Long points out, using the category of indifferents—and, within this category, the possibility of regarding some indifferents as preferred and others as dispreferred—is not denying them all positive value.” The term was meant to capture our attention, as our default is to imagine we can trade virtue for other goods, but putting it “somewhat hyperbolically” is a way to emphasize that “a human being’s fundamental identity and well-being are grounded in her only inalienable possession—herself, her mind or moral purpose” (2002: 24). We are not imagined to ever be indifferent to “health or wealth,” indeed we will always need such indifferents for virtue. And as Annas also puts it, “because the Stoics limit ‘good’ to virtue, they cannot say that some of these [indifferents] are greater goods than others; rather they have to introduce the terminology of calling some ‘preferred indifferents’ and others ‘dispreferred indifferents,’” and the kind of value they have “matters when one is selecting among them” (1993: 167). Some definitions of virtue suggest it is selecting between indifferents properly. So “indifferents” are not meant to suggest we be “totally indifferent to health or wealth” (Long 2002: 245). The term is instead supposed to remind us of the priority of our moral agency. This, again, reminds us that we must turn to ethics to offer fully normative explanations of what we pursue.⁷

⁷ “Virtue has a value which is different in kind from the value of other things like health and wealth. This difference is expressed in several ways, one of which is the way that the Stoics invent a whole new terminology, saying that virtue is chosen whereas the indifferents are only selected, and another of which is the way that these things are said to be indifferents from the viewpoint of virtue, even though we have ‘impulses’ for some of them and ‘counter-impulses’ away from others. That is, we go for some things like health, and avoid others, like illness. Why do we do so, especially if, not being virtue, they do not serve to make up the happy life? The answer is that we go for these things because it is *natural* for us to do so. There would be something wrong with a human who consistently shunned good health and sought sickness” (Annas 1993: 167).

And most of us are aware of this priority. My most colorful example is always the “Texas Cheerleader mom” who took out a hit on her daughter’s rival. We say we would do anything for our children and want the best for them, but we do not mean it in the way this mom did. We mean that we will do for our children only what we can within the constraints imposed on us by morality. We “love” our children to the extent we can as ethical agents. This is a way of saying our daughters are “indifferents” to us: they cannot take priority over our need to be moral.

I can, at this point, imagine objections. But let’s focus not on a wholesale endorsement of Stoicism, and instead on how the concept of indifferents is complex enough to track the difference between economic goods (whether regarded as objective or subjective) and ethics itself. What I propose is that we begin to regard the study of economics, in all of its variety, as the study of “moral indifferents.” Again, this would not mean that economic goods are not crucial to our survival and well-being, but just that they are not the same as virtue. The use of the term “indifferents” gives us two yardsticks: one concerning virtue and the other concerning the value of the indifferents at hand. This is an accessible way for the “ethical” and “economic” levels of analysis to be acknowledged by agents making decisions.

Also, the reminder that economic goods are indifferents is a way to point out that we pursue indifferents collectively. In his *Meditations* (circa AD 161–180), Marcus Aurelius wrote that we must tend the garden of indifferents even if we do not intend to partake ourselves. Though virtue cannot be scarce or used instrumentally, indifferents certainly are and can. Introducing the idea that economic goods are “indifferents” introduces the role of justice in a way that “maximize surplus” certainly does not. Economic benefits are indefensible unless we defend other human rights alongside them. As an example, McCloskey points out that the granting of a “novel liberty and dignity” slowly extended to “all commoners” was what made our current economic climate possible (2016: 413). These are not going to be generated or easily attached to a strategy of maximizing surplus, other than wishfully or *post hoc*. Regarding economic goods as “indifferents,” calling them such, is an active way of holding a place for the ethics that justifies their pursuit, collectively and personally.

2.5 CONCLUSION

To sum up, the three questions with which we started—what must an agent know, how might economics be relevant to virtue, and when can ethics trump economic analysis—can be answered as follows. A virtuous agent must have the right norms about economic “routes” to general well-being so that she is aiming properly at general well-being. Market norms, even if generally applicable, are defeasible. Ethics can trump economic norms whenever a person is justified in seeing other elements of well-being (such as justice or concern for general humanity) as having more significance than the long-term goals met by a functioning market. You may give away your factory to your employees in

your will, even though an economic analysis might show you would maximize surplus by putting it up for sale. You may give your business to Cratchit. You may allow children to pick out a free gift in your store. You may donate your earnings to others. You can bequeath the factory to your employees (Tims 2010). You can decide to pay teaching adjuncts what students expect them to be paid. You just need to be able to defend the good you are doing, in your own words. This is a quite reasonable bar, but again, powerful and revisionary if put to use.

REFERENCES

- Ainslie, George. 2013. "Money as MacGuffin: A Factor in Gambling and Other Process Addictions." In Neil Levy (ed.), *Addiction and Self Control: Perspectives from Philosophy, Psychology, and Neuroscience* (Oxford: Oxford University Press), pp. 16–37.
- Alvey, James. 2011. *A Short History of Ethics and Economics*. Northampton, MA: Edward Elgar Publishing.
- Anderson, Elizabeth. 1995. *Value in Economics and Ethics*. Cambridge, MA: Harvard University Press.
- Anderson, Elizabeth. 2015. "Liberty, Equality, and Private Government." The Tanner Lectures in Human Values, Princeton University. Available at <https://tannerlectures.utah.edu/Anderson%20manuscript.pdf>.
- Angner, Erik, and George Loewenstein. 2012. "Behavioral Economics." In Uskali Mäki (ed.), *Handbook of the Philosophy of Science: Philosophy of Economics* (Amsterdam: Elsevier), pp. 641–90.
- Annas, Julia. 1993. *The Morality of Happiness*. Oxford: Oxford University Press.
- Annas, Julia. 2001. "Introduction." In Cicero, *On Moral Ends*. Julia Annas (ed.) (Cambridge: Cambridge University Press), pp. ix–xxvii.
- Annas, Julia. 2011. *Intelligent Virtue*. Oxford: Oxford University Press.
- Aurelius, Marcus. 161–80. *Meditations, Books 1–6*. Christopher Gill (trans.). Oxford: Oxford University Press (2013 edition).
- Backhouse, Roger E., and Beatrice Cherrier. 2017. "The Age of the Applied Economist: The Transformation of Economics Since the 1970s." *History of Political Economy* 49 (supp): 1–37.
- Baker, Jennifer A. 2009. "Virtue and Behavior." *Review of Social Economy* 67: 3–24.
- Baker, Jennifer A. 2013. "Virtue Ethics and Practical Guidance." *Social Philosophy and Policy* 30: 297–313.
- Baker, Jennifer A. 2015. "Visible Hands: Ethics and the Market." In Charlotte Thomas (ed.), *Of Sympathy and Selfishness: The Moral and Political Philosophy of Adam Smith* (Mercer, GA: Mercer University Press), pp. 78–96.
- Baurmann, Michael, and Geoffrey Brennan. 2016. "On Virtue Economics." In Jennifer A. Baker and Mark D. White (eds.), *Economics and the Virtues: Building a New Moral Foundation* (Oxford: Oxford University Press), pp. 119–40.
- Becker, Gary. 1976. *The Economic Approach to Human Behavior*. Chicago: University of Chicago Press.
- Becker, Lawrence. 2017. *A New Stoicism*. Princeton, NJ: Princeton University Press.
- Bruni, Luigino, and Robert Sugden. 2013. "Reclaiming Virtue Ethics for Economics." *Journal of Economic Perspectives* 27: 141–64.

- Chetty, Raj. 2015. "Behavioral Economics and Public Policy: A Pragmatic Perspective." *American Economic Review Papers and Proceedings* 105: 1–33.
- Cowen, Tyler. 2013. "Is Christmas Efficient?" *Marginal Revolution*, December 23, at <http://marginalrevolution.com/marginalrevolution/2013/12/are-seasonal-business-cycles-inefficient.html>.
- Dasgupta, Partha. 2003. "Modern Economics and Its Critics." In Uskali Mäki (ed.), *Fact and Fiction in Economics: Models, Realism and Social Construction* (Cambridge: Cambridge University Press), pp. 57–89.
- Dickens, Charles. 1843. *A Christmas Carol*. London: Chapman & Hall. Available at <https://www.gutenberg.org/files/46/46-h/46-h.htm>.
- Fehr, Ernst, Urs Fischbacher, and Simon Gächter. 2002. "Strong Reciprocity, Human Cooperation, and the Enforcement of Social Norms." *Human Nature* 13: 1–25.
- Fortenbaugh, William. 2002. *On Stoic and Peripatetic Ethics: The Work of Arius Didymus*. Rutgers, NJ: Rutgers Classics.
- Frey, Bruno S. 1997. *Not Just for the Money: An Economic Theory of Personal Motivation*. Cheltenham, UK: Edward Elgar Publishing.
- Gintis, Herbert. 2000. "Strong Reciprocity and Human Sociality." *Journal of Theoretical Biology* 205: 169–79.
- Görgemanns, Herwig. 1983. "Oikeiōsis in Arius Didymus." In William W. Fortenbaugh (ed.), *On Stoic and Peripatetic Ethics: The Work of Arius Didymus* (New Brunswick, NJ: Transaction Books), pp. 165–89.
- Griswold, Charles. 1999. *Adam Smith and the Virtues of Enlightenment*. Cambridge: Cambridge University Press.
- Hanley, Ryan Patrick. 2015. "On Vernon's Smith." In Charlotte Thomas (ed.), *Of Sympathy and Selfishness: The Moral and Political Philosophy of Adam Smith* (Mercer, GA: Mercer University Press), pp. 97–124.
- Harford, Tim. 2015. "In Praise of Scrooge." *Financial Times*, December 18, available at <http://timharford.com/2015/12/in-praise-of-scrooge/>.
- Harsh, Mel. 2010. "Confessions of Former Debt Collectors." CNN.com, at http://money.cnn.com/galleries/2010/news/1007/gallery.debt_collectors/.
- Hausman, Daniel, and Michael S. McPherson. 1996. *Economic Analysis and Moral Philosophy*. Cambridge: Cambridge University Press.
- Hayek, Friedrich A. 1978. *Studies in Philosophy, Politics, Economics and the History of Ideas*. Chicago: University of Chicago Press.
- Henrich, Joseph, Robert Boyd, Samuel Bowles, Colin Camerer, Ernst Fehr, Herb Gintis, et al. 2005. "'Economic Man' in Cross-Cultural Perspective: Behavioral Experiments in 15 Small-Scale Societies." *Behavioral and Brain Sciences* 28: 795–855.
- Inwood, Brad. 2005. *Reading Seneca: Stoic Philosophy at Rome*. Oxford: Oxford University Press.
- Johnson, Brian. 2014. *The Role Ethics of Epictetus*. Lanham, MD: Lexington Books.
- Landsburg, Steven E. 2004. "What I Like about Scrooge." *Slate*, December 9, at http://www.slate.com/articles/life/holidays/2004/12/what_i_like_about_scrooge.html.
- Long, A. A. 2002. *Epictetus: A Stoic and Socratic Guide to Life*. Oxford: Oxford University Press.
- Long, A. A. 2006. *From Epicurus to Epictetus: Studies in Hellenistic and Roman Philosophy*. Oxford: Oxford University Press.
- Lutz, Mark. 1999. *Economics for the Common Good: Two Centuries of Economic Thought in the Humanistic Tradition*. London: Routledge.

- Mandeville, Bernard. 1724. "The Fable of the Bees: Or, Private Vices, Publick Benefits: With an Essay on Charity and Charity-Schools." In Henry C. Clark (ed.), *Commerce Culture and Liberty: Readings on Capitalism Before Adam Smith* (Indianapolis: Liberty Fund, 2003), pp. 203–39.
- McCloskey, Deirdre. 2006. *The Bourgeois Virtues: Ethics for an Age of Commerce*. Chicago: University of Chicago Press.
- McCloskey, Deirdre. 2016. *Bourgeois Equality: How Ideas, Not Capital or Institutions, Enriched the World*. Chicago: University of Chicago Press.
- Meikle, Scott. 1995. *Aristotle's Economic Thought*. Oxford: Oxford University Press.
- Munger, Michael, and Daniel Russell. 2017. "Can Profit-Seekers Be Virtuous?" In Eugene Heath, Byron Kaldis, and Alexei Marcoux (eds.), *The Routledge Companion to Business Ethics* (London: Routledge), pp. 113–30.
- Natali, Carlo. 1995. "Oikonomia in Hellenistic Greek Thought." In Andre Laks and Malcolm Schofield (eds.), *Justice and Generosity: Studies in Hellenistic Social and Political Philosophy* (Cambridge: Cambridge University Press), pp. 95–128.
- Natali, Carlo. 2003. "Loikonomikos nella tradizione stoica." In Carlos Levy et al. (eds.), *Art et ratio: Sciences, arts et métiers dans la philosophie hellénistique et romaine* (Paris-Bruxelles: Latomus), pp. 75–88.
- Nussbaum, Martha, C. 1988. "Non-Relative Virtues: An Aristotelian Approach." *Midwest Studies in Philosophy* 13: 32–53.
- Phelps, Edmund. 2013. *Mass Flourishing: How Grassroots Innovation Created Jobs, Challenge, and Change*. Princeton, NJ: Princeton University Press.
- Rhonheimer, Martin. 2015. "The Ethics of the Market Economy: A Critical Appraisal of Ludwig von Mises' Utilitarian Interpretation." In Konrad Hummler and Alberto Mingardi (eds.), *Europe, Switzerland and the Future of Freedom: Essays in Honour of Tito Tettamanti* (Torino, Italy: IBL Libri), pp. 353–77.
- Robbins, Lionel. 1932. *An Essay on the Nature and Significance of Economic Science*. London: Macmillan.
- Ross, Don. 2014. *Philosophy of Economics*. London: Palgrave Macmillan.
- Russell, Daniel. 2009. *Practical Intelligence and the Virtues*. Oxford: Oxford University Press.
- Sandel, Michael. 2012a. "How Markets Crowd Out Morals." *Boston Review*, May 1, at <http://bostonreview.net/forum-sandel-markets-morals>.
- Sandel, Michael. 2012b. *What Money Can't Buy: The Moral Limits of Markets*. New York: Farrar, Straus and Giroux.
- Schliesser, Eric. 2016. "The Separation of Economics from Virtue." In Jennifer A. Baker and Mark D. White (eds.), *Economics and the Virtues: Building a New Moral Foundation* (Oxford: Oxford University Press), pp. 141–64.
- Sen, Amartya. 1987. *On Ethics and Economics*. Oxford: Basil Blackwell.
- Sharpe, Matthew. 2014. "Stoic Virtue Ethics." In Stan van Hooft et al. (eds.), *The Handbook of Virtue Ethics* (London: Routledge), pp. 28–41.
- Simler, Kevin and Robin Hanson. 2018. *The Elephant in the Brain: Hidden Motives in Everyday Life*. Oxford: Oxford University Press.
- Smith, Vernon. 2009. *Rationality in Economics: Constructivist and Ecological Forms*. Cambridge: Cambridge University Press.
- Stigler, George. 1975. "Smith's Travels on the Ship of State." In Andrew S. Skinner and Thomas Wilson (eds.), *Essays on Adam Smith* (Oxford: Clarendon Press), pp. 237–46.

- Swanton, Christine. 2016. "Virtues of Productivity Versus Technician Rationality." In Jennifer A. Baker and Mark D. White (eds.), *Economics and the Virtues: Building a New Moral Foundation* (Oxford: Oxford University Press), pp. 185–201.
- Tims, Dana. 2010. "On 81st Birthday, Oregon Man Gives Company to Employees." *Seattle Times*, February 17, at <https://www.seattletimes.com/business/on-81st-birthday-oregon-man-gives-company-to-employees/>.
- Waldfoegel, Joel. 2009. "You Shouldn't Have: The Economic Argument for Never Giving Another Gift." *Slate*, December 8, at http://www.slate.com/articles/business/the_dismal_science/2009/12/you_shouldnt_have.html.

CHAPTER 3

WITH ALL DUE RESPECT

A Kantian Approach to Economics

MARK D. WHITE

IN what is perhaps the best-known (albeit controversial) definition of economics, Lionel Robbins (1932) focused on the allocation of scarce resources to competing ends.¹ This allocation can be conducted in any number of ways, but for most economists since the early twentieth century, the techniques emphasized and promoted as the tools of the discipline—such as preference-satisfaction or utility-maximization, cost-benefit analysis, and social welfare functions—have been *consequentialist* in nature, judging choices and allocations according to their results. More precisely, these tools have been vaguely *utilitarian*: the results of these choices are normally stated in terms of happiness or well-being, whether individual or aggregated. Consumers make choices among goods and services to get satisfaction or utility; firms make decisions regarding price, output, and personnel to earn profit; governments make policy choices using proxies for social welfare (such as gross domestic product or income per capita) and other measures related to it (such as unemployment).

These models are easily translated into mathematics to derive theoretical implications that can be tested empirically in a straightforward manner, but they also limit the range of behavior that economics can describe, understand, and predict. Not all decisions are made on utilitarian grounds, nor are all decision-makers utilitarians. Furthermore, not all economists are utilitarians, and some may not realize that their modeling tools assume and imply utilitarian behavior, much less acknowledge the limits on their descriptive and predictive power that this orientation involves. As John Hicks wrote, “If one is a utilitarian in philosophy, one has a perfect right to be a utilitarian in one’s economics. But if one is not...one also has the right to an economics free from utilitarian assumptions” (1939: 18). It is not only practicing economists, but the consumers of economics as well, who deserve an economics that incorporates other ethical perspectives alongside—and not necessarily to the exclusion of—utilitarianism.

¹ For the history and criticisms of Robbins’ definition, see Backhouse and Medema (2009).

To this end, in this chapter I will explore the possibility of an economics that includes the deontological moral philosophy of Immanuel Kant, which emphasizes principle and rights above considerations of preference and welfare. After introducing the basic elements of Kant's moral theory—autonomy, dignity, and duties both perfect and imperfect—I will suggest one way that Kantian duty can be incorporated into the standard model of constrained preference satisfaction, allowing the economic agent to make principled decisions alongside prudential ones. Then I will survey several areas of economics in which a Kantian approach can make an important contribution, including instances of individual choice in which ethical factors are likely to be particularly important, as well as policy issues which are inherently ethical in nature, including the way the state intervenes in market activity.

3.1 DEONTOLOGY AND KANTIAN ETHICS

Immanuel Kant's ethics can be classified broadly as a type of *deontology*, a school of ethics that prioritizes inherent properties of acts over their consequences, a distinction reflected in the title of deontologist W.D. Ross's 1930 book, *The Right and the Good*.² Deontologists would regard acts as wrong because they violate some moral principle or right, whereas consequentialists would make such a judgment based on the likely outcomes of the acts.

Although deontologists and consequentialists often agree on basic moral judgments regarding lying, theft, and murder, they do split on cases in which good outcomes are possible only if deontological principles are violated. The most well-known (if infamous) case of this is the *trolley problem*, in which an out-of-control trolley car is hurtling toward a broken section of track that, if struck, will result in the death of all five passengers aboard (Foot 1967; Thomson 1976). A bystander, observing this impending calamity, has the opportunity to pull a switch and divert the car to another track, which will save the five trolley passengers but at the cost of the death of a person sitting blissfully unaware on the other track. The consequentialist would likely argue that the bystander should pull the switch to save five lives at the expense of one, whereas the deontologist would question whether the bystander can or should directly act to kill the one person even to save the five. Other variants of this basic dilemma trigger different intuitions and reactions: one variant of the trolley problem has the bystander push a heavy man in front of the trolley to stop it (rendering the action more directly tied to his death), and a different scenario altogether imagines a surgeon treating five urgent transplant patients who can all be saved if she harvests the organs of her healthy receptionist. What these thought experiments share is

² For more on the meaning and varieties of deontology, see Gaus (2001a; 2001b), as well as the chapter by Zamir and Medina in this volume.

a focus on doing a presumptive moral wrong for the greater good, which is key to the conflict between deontology and consequentialism.³

Kant's deontological ethics derives its moral principles from a rich, humanistic background.⁴ The heart of his moral philosophy lies in *autonomy*, the capacity and responsibility of all rational beings (or persons) to determine their own wills and make decisions for the right (moral) reasons even when such choices go against their own preferences or desires: "Autonomy of the will is the property that the will has of being a law to itself" (Kant 1785: 441). This stands as a challenge to the implicit assumption of *homo economicus*, who is beholden to satisfy his preferences as much as possible (within his constraints) and has no choice otherwise (and no will at all). The Kantian sense of autonomy implies that we can recognize reasons for action other than our own preferences and act on them when morality demands it.⁵

In turn, autonomy imbues persons with *dignity*, an "unconditional and incomparable worth" that entitles them to respect and consideration from others (1785: 436). It is this dignity that distinguishes persons from things: "whatever has a price can be replaced by something else as its equivalent . . . whatever is above all price, and therefore admits of no equivalent, has a dignity" (1785: 434). A thing is instrumental, merely a means to an end, with its value based on its usefulness, whereas persons are "ends-in-themselves" with an intrinsic, incomparable value. Kant's dignity also stands as a strong affirmation of moral equality; as Thomas E. Hill writes, "the root idea of dignity is simply that virtually everyone, regardless of social station, talents, accomplishments, or moral record, should be regarded with respect as a human being" (1991: 170).

Autonomy and dignity are put into practice through the *categorical imperative*, Kant's formalization of "the moral law," which is commonly expressed in three formulae that can be used to test plans of action, or *maxims*, for their ethical permissibility. The most familiar one is the Formula of Universal Law: "act only according to that maxim whereby you can at the same time will that it should become a universal law" (Kant 1785: 421). This formula maintains that a maxim is only allowed if everyone can be allowed to adopt the same maxim without contradiction, in acknowledgment of the dignity of all. This moral equality is more obviously apparent in the Formula of Respect: "act in such a way that you treat humanity, whether in your own person or in the person of another, always at the same time as an end and never simply as a means" (1785: 429). This formula demands that all persons (including oneself) be treated as ends in themselves, which is usually taken, at minimum, to forbid actions involving deceit or coercion, but also to encourage kindness and altruism towards all. The final version, the Formula of the Kingdom of Ends—"every rational being must so act as if he were through his maxim always a legislating member of the universal kingdom of ends" (1785: 438)—is less often

³ The trolley problem has enjoyed renewed "popularity" recently, especially in light of the development of autonomous cars which must be programmed to deal with similar situations. For a sampling of the popular and academic discussions, see Cathcart (2013), Edmonds (2014), and Kamm (2016).

⁴ The bulk of Kant's ethics can be found in Kant (1785, 1797). For overviews, see Sullivan (1989, 1994).

⁵ For more on Kantian autonomy, see Hill (1989) and Irwin (2004).

used to evaluate maxims, instead representing the ideal of self-legislation and the moral community that can result from it.⁶

The details of the categorical imperative are not as important for our purposes as are the results of applying it. When a maxim is rejected by the categorical imperative, this rejection implies a *duty* not to act on such a maxim. For example, a maxim of lying for one's own benefit would be rejected by the Formula of Universal Law, because opportunistic lying, when universalized, would erode trust and make such lies ineffective; and the Formula of Respect, because lying uses other persons simply as a means to the liar's ends without treating them as ends-in-themselves. This rejection creates a duty not to lie for one's own benefit, which is ultimately grounded in the equal dignity (and autonomy) of all persons. By similar logic, a maxim of selfishness—being indifferent to the suffering of others—is also rejected, most clearly by the Formula of Respect (because it fails to consider the needs of others), resulting in a duty of beneficence.⁷

These two examples correspond to the two types of Kantian duties: perfect and imperfect. *Perfect duties* are those that admit of no exceptions for the sake of preference or desire (even if other-regarding). These are usually negative in nature, such as duties not to lie, steal, and kill, all of which command us not to engage in actions that fail the categorical imperative test. *Imperfect duties* do not prescribe certain action or inaction, but rather require attitudes that should be expressed in action when possible. The best example of an imperfect duty is that of beneficence, which we should practice whenever we can, taking into account our other duties (perfect and imperfect), as well as (to some extent) our own preferences, desires, and well-being, which fall under our duty of beneficence to ourselves (as persons equally deserving of respect and consideration).⁸ Although most imperfect duties are positive in nature, encouraging a certain kind of behavior rather than proscribing it, they also result from rejection of a contrary maxim by the categorical imperative test: as described above, the duty of beneficence is actually a duty not to be indifferent to the suffering of others.⁹

Several aspects of duties are important to note even at this broad level of description. First, although persons should always act *according* to duty, it is more important for their moral characters that they do so for the *sake* of duty. Kant gives the example of a businessman who refuses to cheat his customers, which is dutiful, but may do so merely to ensure a good reputation and increase his sales, rather than for the sake of following the moral law (1785: 397). Kant is realistic regarding our often mixed motivations for action—his ethical businessman may have both ethical and prudential reasons for being

⁶ For more on the categorical imperative, see Sullivan (1989: chs 2–6; 1994: Part III) and Paton (1947).

⁷ This maxim is also rejected by a different version of the universalization formula, the Formula of the Law of Nature—"Act as if the maxim of your action were to become through your will a universal law of nature" (Kant 1785: 421)—under the reasoning that we could not rationally will to live in a world of universalized indifference to each other.

⁸ The flexibility of imperfect duties explains why they are also called *wide* duties, as opposed to *narrow* perfect duties. Perfect duties are also called duties of *action* whereas imperfect duties are duties of *ends*, given their disparate focus.

⁹ For more on perfect and imperfect duties, see Gregor (1963: ch. 7).

honest—but warns against relying on self-interest for ethical behavior because it may not always lead us in the right direction.

Second, the derivation of duties from the categorical imperative is fraught with ambiguity. It is unclear exactly how maxims are to be crafted and in what way they must be inconsistent when universalized, or treat humanity simply as a means, to result in a duty. Herman (1989) shows that the Formula of Universal Law, which Kant proposed as the best formula for judging maxims, fails to generate a duty against occasional murder—which stands as a reasonable litmus test for any ethical system—whereas the Formula of Respect easily and definitively condemns it. This is less of a criticism of the categorical imperative than the way in which it is used, and demonstrates that it should be regarded as a way to think about morality and our duties rather than a mechanistic algorithm for generating them. The spirit of dignity and respect that underlies the categorical imperative is more important than whether a certain formula rejects a particular maxim depending on how it is worded.

Third, there are inevitably times when one's duties may seem to conflict, a situation which Kant acknowledged but about which he provided little guidance. In such cases, our one and only true duty is determined by which action (or inaction) has the "stronger ground of obligation" (1797: 224), on which he provided no elaboration. One interpretation is that he left such cases up to *judgment*, which Kant described as "a peculiar talent which can be practiced only, and cannot be taught. It is the specific quality of so-called mother-wit; and its lack no school can make good" (1781/1787: A133/B172). Judgment is necessary not only in cases of conflicting duties, but also in the derivation and application of duties, given the ambiguity in using the categorical imperative.

This need for judgment to guide the implementation of duty is not a flaw but a feature. The categorical imperative is best thought of as a quick-and-easy way to apply the moral law to general actions, after which we apply the vague duties it generates to specific decision-making circumstances in a way that incorporates real-world context. As Onora O'Neill explains, Kant "never assumes agents can move from principles of duty, or from other principles of action, to selecting a highly specific act in particular circumstances without any process of judgment" (2004: 104). Instead, as Kant himself acknowledged, duties require "a power of judgment sharpened by experience, partly in order to distinguish in what cases they are applicable, and partly to gain for them access to the human will as well as influence for putting them into practice" (1785: 389), as well as knowledge of psychology or, as he called it, "anthropology in order to be *applied* to humans" (1785: 412).¹⁰

The role that judgment plays in humanizing Kant's moral theory helps to counter common misperceptions that it is extreme and demanding. It also helps to remember that all Kantian duties are technically negative, steering us away from actions that violate the categorical imperative—do not lie, do not kill, do not be indifferent to the suffering of others—but without dictating what we should do instead. For example, we must not lie,

¹⁰ "A metaphysics of morals cannot dispense with principles of application, and we shall often have to take as our object the particular nature of human beings, which is cognized only by experience, in order to show in it what can be inferred from universal moral principles" (Kant 1797: 216–17).

but this does not imply that we must be completely forthright; as long as we do not deceive, we can prevaricate, change the question, or simply refuse to answer. We must not be completely indifferent to others' pain, but we do not have to devote our entire lives to the service of others; as Kant wrote, "How far should one expend one's resources in practicing beneficence? Surely not to the extent that he himself would finally come to need the beneficence of others" (1797: 454). Both perfect and imperfect duties leave significant latitude for action within constraints implied by the moral law; while Kant certainly encouraged action above and beyond the strict requirements of duty, he stopped short of requiring it. This strong role for judgment in balancing duties, applying them to real-world circumstances with context, and choosing actions within the constraints of duty, makes his ethics far less demanding than perfectionist versions of utilitarianism, and closer to classical virtue ethics, with its emphasis on judgment and character.¹¹

3.2 KANTIAN DUTIES IN MODELS OF ECONOMIC CHOICE

3.2.1 Perfect and Imperfect Duties

Having described Kant's perfect and imperfect duties, we can now suggest a way to work them into economic models of choice, inspired by and building on previous work on commitment by Sen (1977) and Minkler (2008).¹² As the traditional economic model is normally presented, agents have a set of well-ordered preferences that define their utility or "well-being" but are not derived from it. In mainstream economics, to maintain generality, no content or psychological basis is imposed on preferences; as a result, an agent's preference ordering may be based on narrow self-interest, concern for friends and family, altruism towards one's community, or negative feelings such as animus toward certain groups. Agents make choices to satisfy their preferences maximally given their constraints, which are normally understood to be based on resources, most important among them money and time. These constraints serve to limit the choices agents can make and force them to choose from the available options those that will satisfy the most or highest preferences.

Both an agent's preferences and constraints are normally understood to be exogenously given, which may be a reasonable heuristic assumption at a point in time, but over a length of time both can be influenced by agents themselves as well as outside forces. This is easier to see with reference to constraints: for example, an agent's future wealth can be

¹¹ There is an enormous literature on Kant and virtue ethics: see, for instance, Engstrom and Whiting (1996), Sherman (1997), Betzler (2008), Jost and Wuerth (2011), and Cureton and Hill (2015). (I discuss this link in the context of economics in White 2016a.)

¹² The following is based on White (2011a: ch. 1).

increased by making certain decisions in the present with regard to work and savings. However, preferences can be developed and nurtured by agents over time as well: a person who meets a fan of jazz may start to appreciate it despite an initial disdain, and may even choose to cultivate such a desire independently.¹³

If we understand preferences and constraints to be (to some extent) subject to conscious control, we can imagine agents choosing to follow Kantian duties autonomously by incorporating them into their decision-making framework. As it happens, perfect and imperfect duties correspond closely to constraints and preferences, respectively. Perfect duties, being mostly negative in nature, represent constraints on action, limiting the choices agents can make to satisfy their preferences. Just as agents may not spend more money or time than they have on a certain activity, agents that accept Kantian duty cannot lie, steal, or hurt other people for their own ends. Resource constraints may be less voluntary, at least in the short run, while duty must be freely acknowledged and accepted to imply true moral character. If agents make a choice to follow the moral law and make ethical choices autonomously, then perfect duties can be as binding on them as resource constraints are. Ironically, resource constraints can be violated if moral constraints are dismissed, such as when people embezzle money from work or cheat on their taxes. The fact that most agents in economic models are assumed to follow basic principles of honesty in market behavior implies that such moral concepts are implicit in economics, not antithetical to it.

Likewise, imperfect duties can fit among an agent's preferences. Because imperfect duties are attitudes or ends that should be followed when possible, they represent goals that sit alongside an agent's self-interested and other-regarding preferences, so the agent chooses amongst all of them when making choices within constraints. When a person decides how to divide her paycheck between spending and saving, she may also devote a portion of it to charity out of her duty of beneficence. Because imperfect duties are wide and flexible, she might not be able to donate as much of her income when her other expenses are higher, or even when she wants to splurge on herself on occasion, but should donate more when she is able. In this sense, imperfect duties resemble other choice options, the salient difference being that they are intentionally cultivated and pursued, whereas most preferences are the result of past experiences and actions. Again, agents can certainly choose to ignore their imperfect duties and make purely self-interested choices—much as they can choose to ignore the constraints of perfect duty—but autonomy implies that they *can* choose to follow duty even when it stands in the way of satisfying their other preferences.

3.2.2 Judgment and Willpower

Although the agent is free to include perfect and imperfect duties into her constraints and preferences, there is still the problem of deciding what her duties *are* as well as

¹³ For a survey of theories of preference change, see Grüne-Yanoff and Hansson (2009).

exactly *how* to include them in her decision-making process, which brings us back to the topic of judgment. Perfect duties are simpler in this regard: as described above, adopting duties not to lie, steal, and kill works similarly to acknowledging resource constraints, and can generally be observed simultaneously, all serving to narrow down the set of possible actions. Imperfect duties, however, are more complicated: merely acknowledging a duty of beneficence does not tell the agent where to rank it among her other preferences. She may rank donating money to charity above buying new clothes, but not if those clothes are necessary for a job interview. There is no hard and fast rule that dictates how various acts of beneficence are to be ranked amongst actions based on preference; in this ubiquitous case, judgment is indispensable.

This need for judgment in economic decision-making is even more obvious when duties conflict, particularly when principled action becomes very costly, imperiling other morally important ends. Looked at simply, it seems there can be no exception to perfect duties against stealing or killing, yet there can always be a more important principle at stake, such as feeding your starving family or protecting a life.¹⁴ Sometimes that endangered principle can be put in the context of resources: keeping an ill-considered promise to give your shiftless brother \$10,000 to start yet another foolish business venture prevents you from using that money to do actual good in the world. Standing by principle is noble but can be self-defeating if it comes at the cost of other valuable principles or goals, and judgment is necessary to determine when that line is crossed. This analysis is similar to *threshold deontology*, which maintains that deontological rules are to be followed until they reach a certain level of costs, after which consequentialist considerations take over (Moore 1989). Even here, however, judgment is necessary to determine where the threshold lies and how strictly the threshold condition is to be followed (White 2011b).¹⁵ This way of looking at principled behavior also acknowledges that principles are subject to opportunity cost that can influence their practice without denying or violating their unique status as principles (White 2015a).

Even if we assume fine judgment on the part of economic actors, actual persons in the real world are imperfect and may not be willing or able to obey its dictates to the letter. Kant was as aware of this as anybody, maintaining that only divine beings can be perfectly moral, while human beings are physical beings who must struggle against their natural desires and inclinations to follow the moral law: “The perfect fit of the will to moral law is holiness, which is a perfection of which no rational being in the world of sense is at any time capable” (1788: 122). In other words, it takes fortitude, strength of character, or virtue (as Kant used the term) to resist our impulses and follow our duty as determined by our judgment: “For while the capacity to overcome all opposing sensible impulses can and must be simply *presupposed* in man on account of his freedom, yet this capacity as *strength* is something he must acquire” (1797: 397). Even the strongest among us will fail, but less often—or with respect to less weighty duties—than those with less strength.

¹⁴ Kant (1799) infamously wrote that there is no exception to the duty not to lie, but his own position of judgment and conflicts of obligation would seem to counter this; see White 2015a: 231–2.

¹⁵ For more on threshold deontology, see the chapter by Zamir and Medina in this volume.

Furthermore, Kant argued that this strength can be bolstered by use—“the way to acquire [strength] is to enhance the moral *incentive* (the thought of the law), both by contemplating the dignity of the pure rational law in us and by *practicing virtue*” (1797)—similar to psychologists’ work on willpower as a muscle (Baumeister and Heatherton 1996).¹⁶

This Kantian insight can also be incorporated into economic models of choice to provide an alternative account of weakness of will. In mainstream and behavioral economic models of weakness of will, agents’ preferences, or the way they compare and process them, are modified to generate decisions that favor short-term rewards over long-term interests.¹⁷ Despite the insights of this work, they can explain only why agents display weakness of will, not why they are able to *resist* such temptations—why people manage to stick to diets, stay sober, or avoid procrastination. In other words, they are fatalistically deterministic in the way they map preferences and constraints onto choices with no actual choice involved. It is very difficult to model weakness of will in economics because economic models leave no room for a meaningful conception of the will.

One way to incorporate will (and weakness thereof) into economic models is to explicitly portray the struggle agents face between two paths, the moral and the prudential (White 2011a, ch. 2). For example, one path may have an agent working diligently on projects to finish them when promised while the other has him surfing the web; alternatively, one path may have an agent pursuing her goals while maintaining honesty while the other has her lying when she finds it advantageous. This technique borrows from the literature of multiple utilities or preference rankings (Etzioni 1987; George 2001), but here the specific “metachoice” between them is not a choice at all but an act of will, modeled as a stochastic process that is a function of the agent’s strength, resolve, or willpower. The stronger the agent is, the less often he will procrastinate or she will lie, but either is possible unless the agent is of perfect strength (or never faces an opportunity tempting enough). In such a model, an agent facing clear incentives to violate duties may do so, not simply because the incentives are greater, but because their will was not strong enough (which is not necessarily a function of the incentive). Also, the agent may resist temptation, regardless of the incentives behind it, because it is the right thing to do and their will is aligned with that recognition.

3.3 APPLICATIONS OF KANTIAN-ECONOMIC CHOICE

Although not all economic decision-making situations are ethically loaded to the same degree—deciding how much money to donate to charity is not the same as choosing a brand of cat food—all choices potentially have moral aspects that can be modeled,

¹⁶ For more, see White (2011a: 57–9) and references therein.

¹⁷ For example, see economists’ work on procrastination, including Akerlof (1991) and O’Donoghue and Rabin (2001).

understood, or interpreted more comprehensively using a Kantian framework. This section surveys several examples drawn from traditional economic topics such as consumer and producer behavior as well as less traditional ones such as law and marriage.

There is an extensive and insightful literature in economics regarding altruism, most of which examines the trade-offs between spending that satisfies self-oriented preferences and that which benefits others, normally modeled as other-regarding preferences (which themselves can be based on the consumption of others or their welfare) or different types of utility altogether.¹⁸ A Kantian-economic way of modeling altruism resembles these methods, even if ultimately grounded in duty rather than preference or utility. Altruism (or beneficence) is an imperfect duty, and as such appears in an agent's decision-making process alongside other preferences to be satisfied as the agent chooses within available resources. As an imperfect duty, an agent's amount of altruistic giving can be optimized along with self-oriented spending (which can be understood to fulfill the agent's duty to herself) in much the same way that it is modeled in mainstream economic models of altruism, the main difference being that such beneficent giving would not be motivated by any sense of utility (Etzioni 1988) or "warm glow" (Andreoni 1990), but would be performed for the sake of duty (although multiple motivations often do exist).

Where a Kantian frame becomes more important is when altruistic giving is accompanied by a promise or commitment. As Sen (1977) recognized, commitment changes the consideration of trade-offs, fixing one aspect of the maximization problem and transforming it into a sort of constraint against which other more flexible choices can be made. If the agent commits to giving a certain amount of money to a charity each month or year, then it becomes a promise or perfect duty—which, while not entirely inviolable or unchangeable, represents a *pro tanto* commitment that is not likely to change in response to marginal changes in the agent's economic circumstances, such as income or the prices of other goods and services.

Not all ethically motivated spending takes the form of charitable giving, though. Many purchases of ordinary consumer goods are made based on moral characteristics or properties, such as those dealing with the environment, treatment of animals, fair trade, or the producers' social and political attitudes. Consumers who feel strongly about such issues are willing to pay more for goods and services that align with their concerns (like other preferred characteristics with less ethical importance), and may avoid products that do not, even if they are significantly cheaper or more easily available. Like the altruistic commitment that restricts marginal trade-offs, such consumer behavior reflects the influence of principle or duty. Such consumers are committed to not supporting products or producers that violate deeply held ideas and are willing to bear significant additional costs to do so, behavior that is not amenable to substitutions in response to small changes in relative price.

Principled behavior can also be seen in producers or firms who incorporate ethical considerations into their decision-making. Under the standard assumption of profit maximization, the firm engages in ethical behavior—such as corporate philanthropy, environmental initiatives, or political action—only if it is the best use of resources to

¹⁸ See, for example, Margolis (1984) and Etzioni (1987).

increase profits. Such “strategic” ethical behavior is rightfully criticized as insincere from a moral point of view (leading to terms such as “greenwashing” and “pinkwashing”), and shifts credit to the firm’s customers and business partners, whose ethical preferences drive the firm’s “ethics” in the same way that consumer preferences drive product design.

This does not mean, however, that a firm cannot behave sincerely according to Kantian motives of acting for the sake of duty. Firms can commit to a certain code of behavior—Google’s overly simplistic “do no evil” comes to mind—that prevents them from taking immoral actions to boost the bottom line. Of course, that code may itself have been chosen to increase long-term profits through crafting an identity that would earn loyalty from customers, but this only shows that the motivations for any behavior can be overdetermined and unverifiable, even by agents themselves (as with Kant’s honest shopkeeper). While cynicism may drive us to doubt any corporate motives, there is no reason to exclude entirely the possibility that firms may foreclose some negative behavior (such as abuse of the environment or harsh treatment of employees) and commit to positive behavior (such as philanthropy or support of social causes) out of the sincere ethical motivations of its leaders or owners.¹⁹

Ethical dimensions are more clearly seen in decisions made outside traditional economic contexts. As exemplified by the work of Gary Becker (1976) and others, economists have applied the techniques of marginal analysis and market equilibrium to topics normally outside the scope of economics, such as discrimination, the family, and the law. While these “incursions” into other fields have yielded some important insights about aspects of choice in these situations that resemble market exchange—and have drawn criticism from within economics (Fine 2002) and, naturally, the other fields themselves—the application of the standard utilitarian model of choice omits other ethical perspectives such as duty, commitment, and rights, that give these areas their distinctly non-economic nature.

I will mention just two such fields here: the economic approach to the family and to law. In the first, economists adapt standard microeconomic models of production and exchange to various aspects of family life, such as the formation (and dissolution) of marriages, the division of resources and responsibilities within them, and decisions regarding childrearing (Becker 1981; Bergstrom 1996). To their credit, researchers in this area do not model family members as strictly self-interested but rather explicitly altruistic towards each other (even if opportunistically so, such as in Becker’s “Rotten Kid” theorem). However, altruism is modeled in the standard economic way, which fails to capture the element of commitment and duty that holds families together—and the absence of which explains why they fall apart—as well as aspects of moral imperfection and weakness of will that may be more prevalent in emotionally fraught situations (White 2015b). On a more basic level, the standard model of economic behavior, even with some element of altruism included, fails to capture the intense and complex loving bonds that characterize such relationships, which often invokes deontology, virtue ethics, or the ethics of care.²⁰

¹⁹ Benefit corporations are one institutionalized example; see Hiller (2013).

²⁰ For more on care ethics, see the chapter in this volume by Knobloch on feminist ethics and economics.

While the economics of the family is treated with skepticism outside economics (and to some extent within it), the economic approach to law, or “law and economics,” has been enormously successful, becoming a significant area of study in the legal academy as well as economics departments (Mercuro and Medema 2006). The economic analysis of every area of the law, from torts and contracts to crime and procedure, has injected the language of marginal analysis and trade-offs into the way legal professionals from attorneys to appellate justices think about how the law affects decision-making and ways that process can be made more efficient and beneficial.²¹

However, there are integral ethical aspects to the law and legal decision-making that the standard economic approach to the law omits. At the most basic level, individuals’ choices to obey or break the law are rarely purely self-interested, reflecting to some extent an obligation to obey the law if not a belief in the moral worth of it. The agent of economics, however, resembles Oliver Wendell Holmes’ “bad man” (1897) who cares for the law only insofar as it affects him and therefore obeys the law only when it is in his interest to do so. Also, the behavior of actors in the legal system, such as attorneys and judges, are also modeled in terms of self-interest, ignoring the elements of duty and obligation present in such roles. The neglect of the ethical aspects of legal behavior compromises the explanatory and predictive value of these models, which in turns weakens the case for legal reform and enforcement. Finally, attempts to explain the scope and the content of the law themselves in mainstream economic terms leave out its inherent ethical content. This is most evident in the economics of crime; as Jules Coleman wrote,

such a theory has no place for the moral sentiments and virtues appropriate to matters of crime and punishment: guilt, shame, remorse, forgiveness, and mercy, to name a few. A purely economic theory of crime can only impoverish rather than enrich our understanding of the nature of crime. (1985: 326)

Crime is not alone in this, however; every area of law, including the private law categories of contract, property, and tort, embodies intrinsic deontological concepts of rights and justice that are neglected by the economic approach to law.²²

3.4 KANTIAN IMPLICATIONS FOR WELFARE ECONOMICS, POLICY, AND MARKETS

The utilitarian basis of economics is even more clearly seen in the normative analysis of individuals’ actions, business activity, and government policymaking. Not only are agents understood to make decisions to satisfy their own preferences and maximize their “utility,” but those decisions are judged, and government policy is made, in order to

²¹ See the chapter by Zamir and Medina in the volume for more on law and economics.

²² For more, see White (2011a: 134–62).

maximize aggregate utility. In this area more than any other, economics can be seen as instrumentalized utilitarianism, whether using theoretical constructs such as social welfare functions, more practical techniques such as Kaldor-Hicks efficiency, or real-world measures such as gross domestic product (and, more recently, gross domestic happiness).

3.4.1 Utilitarianism, Kaldor-Hicks, and Pareto

In general, economists' implicit adoption of utilitarianism opens it up to the same criticisms made against the ethical system itself.²³ Most important for our purposes is that, although utilitarianism treats every person's well-being equally in the sum total—reflecting the moral equality that both utilitarians and Kant emphasize—the process of summing also makes individuals' utilities indistinguishable. John Rawls (1971: 27) famously made the point that utilitarianism “does not take seriously the distinction between persons,” rendering them mere receptacles for fungible utility. As long as the total increases, it matters not whether Pauline's or Peter's utility is increased, or that the well-being of a privileged few is increased and that of everyone else is left the same. Distributional concerns regarding fairness and equality are not accounted for in classical utilitarianism, nor in most discussions of welfare economics.²⁴

More worrisome on Kantian grounds, however, is that utilitarian summing implies that each individual's utility is interchangeable or substitutable. Not only will an increase in either Pauline's or Peter's utility increase the total, but a certain increase in Pauline's utility can justify a smaller decrease in that of Peter. To be fair, this is not a problem *per se*. For instance, if initial allotments of a scarce resource are being made with no pre-existing entitlements or claims to equal shares, and there is good reason to give more to Pauline than to Paul, Paul has not lost anything he “had” in any meaningful sense. The crucial point is the existence and nature of pre-existing rights: if Pauline's utility can only be increased by violating Paul's rights, this requires more justification in a Kantian frame than simply an increase in aggregate welfare.

In economics, this trade-off between the well-being of individuals in the name of total utility is known as Kaldor-Hicks efficiency, a policy-oriented version of cost-benefit analysis that mirrors individual decision-making.²⁵ Whereas individuals are free to make trade-offs with their own preferences and welfare when making choices in their own interests, it is another thing entirely for a person or party to trade off one person's utility against another's, especially when doing so violates rights. If policymakers take something from Peter to make Pauline better off, thereby increasing total welfare, Peter has been used merely as a means to the ends of the decision-maker. (Technically, so has

²³ See, for example, Scheffler (1988).

²⁴ On these issues, see the chapters by Broome and Fleurbaey in this volume.

²⁵ James Buchanan (1990: 5–7) warned of the hidden dangers of applying models of individual choice to policy: “This feature allows for a relatively unnoticed transference of analysis from individual to *social* or *collective* choice on the basis on some implicit presumption that collectivities choose analogously to individuals” (6).

Pauline, even though she was unharmed.) This stands in contradiction to Kant's Formula of Respect, which applies to governments as it does for individuals, and demands that no person be used merely as a means without at the same time being treated as an end, which would require consent before any wrongful harm is imposed.

Defenders of Kaldor-Hicks dismiss the consent objection in a number of ways. Most simply, they argue that because the "winners" from a Kaldor-Hicks efficient change gain more than the "losers" lose, the former could compensate the latter and still be ahead, leading to what is called a "potential Pareto improvement," referring to the stricter policy criterion that approves changes that benefit some only if none are harmed. But this is a mathematical truism that simply reframes the issue, without providing a justification for it. As Anthony Kronman wrote,

someone committed to the Kantian idea of individual autonomy would rightly feel that his moral principle had been violated, and it would not make any difference, from his point of view, that compensation could potentially be made even though it was not. *For a Kantian, the Kaldor-Hicks test has no significance.* (1980: 238)

Furthermore, even the Pareto test fails to meet Kantian standards if the judgment of whether someone is hurt or not is made externally (White 2011a: 165–79). For example, when a homeowner whose property is taken by the government under eminent domain is provided a "fair market value" as compensation, this would often be regarded as a Pareto improvement by an outside evaluator. But if the homeowner does not feel the compensation is adequate, then it is not a Pareto improvement, and is not ethical on Kantian grounds despite being Kaldor-Hicks efficient.

In general, consent is necessary to ensure that individuals are not being used merely as means, and the Pareto criterion is no substitute for consent; ideally, they would be the same thing, rendering the Pareto test redundant.²⁶ For Kaldor-Hicks to be ethically relevant from a Kantian point of view, consent of the affected parties would have to be secured, which would presumably require agreeable compensation, preferably incorporated into the policy proposal itself. Economists often complain that the transaction costs of compensation (not to mention consent) are too high, but this should speak against the proposal itself—which, by implication, must have very small net benefits—not the process of arranging compensation and securing consent.²⁷

3.4.2 Rights, Wrongs, and Externalities

Another way to look at a consent requirement in Kaldor-Hicks situations is that it would endow the affected parties with the right of refusal, which leads us to discuss the neglect of any meaningful concepts of rights in utilitarian economics and the corollary absence of any

²⁶ For more on this point, see Coleman (1984).

²⁷ See also the chapter by DeMartino in this volume for more on Kaldor-Hicks efficiency and the resistance to arranging compensation.

notion of wrongfulness (a violation of rights). Although the concept has many dimensions and forms, for our purposes a right can be understood to shield the individual from having to conform to the demands of utilitarianism. This formulation is consistent with Ronald Dworkin's (1977) famous dictum that, in at least some significant cases, rights must "trump" welfare: there must be some nontrivial instances in which a right permits individuals to make choices regardless of their effect on total utility. For example, individuals in a liberal democracy can choose where to live, what job or career to pursue, and who to marry (or not), without being forced to make the choices that would most likely contribute to aggregate well-being.

Utilitarians and economists recognize only those rights that are supported by considerations of welfare, and can be dismissed for the same reason. As Dworkin characterized this view, in economics "the institution of rights, and particular allocations of rights, are justified only insofar as they promote social wealth more effectively than other institutions or allocations" (1980: 243). As such, these "rights" have no power to counter the utilitarian impulse that inspired it, and therefore are not meaningful rights in Dworkin's sense. This is not to say, however, that rights must be absolute to be meaningful; they can be overridden and waived, but only if justified by some aspect of the situation that goes above and beyond simple welfare-maximization. Richard Posner, an advocate of the economic approach to rights, seems to endorse absolute rights, but ends up making Dworkin's point when he writes "the economist recommends the creation of [absolute] rights . . . when the cost of voluntary transactions is low . . . But when transaction costs are prohibitive, the recognition of absolute rights is inefficient" (1983: 70). Again, granting rights solely on the basis on considerations that rights are designed to protect against only defeats their purpose.

In a Kantian context, rights are a corollary of duties—a duty not to steal implies a right not to be stolen from—and, in general, not being treated merely as a means, especially to increasing the general welfare. Not only does this provide us with another way to frame the objection to Kaldor-Hicks efficiency, but it also allows us to critique another concept taken for granted in mainstream economics: externalities. Externalities are effects (usually deleterious) on parties not involved in a transaction, with pollution being the standard example, and are a significant concern in assessments of the welfare effects of market activity, especially because, outside of government intervention, the participants in transactions have no incentive to account for third-party harms.

The standard way to deal with externalities is through Pigouvian taxes, which force the parties creating the harm to internalize the costs they are imposing on others and adjust their behavior accordingly (Pigou 1932; Baumol 1972). Although Pigouvian taxes potentially make the situation efficient from the view of total welfare, they do not eliminate externalities, but merely reduce them to the level where any further reductions in harm would cost the responsible parties more than the harm they would save the third parties. In other words, Pigouvian taxes shift the incidence of harm while lowering it, but they do not recognize any rights of the injured to be free of harm—and implicitly endow the injurer with the right to continue imposing harm (with the incentive to do it at the overall efficient level).

In the work that led to his eponymous theorem, Ronald Coase (1960) emphasized the role played by rights in the efficient solution of externalities. If the injured party has the right to compel the injurer to correct the harm, the injurer will do so in the lowest-cost way, even if that means buying the right to continue the harmful activity from the injured—but only if the injured party agrees, based on their right. If the injurer has the right to continue the activity, then the injured party has the same options: either solve it themselves or buy the right from the injurer, if they agree. Coase's point was that the efficient solution will obtain regardless of who has the right to control the situation, but his analysis also serves to show how important rights (and the voluntary exchange of them) are important to economic and legal transactions.

Where this focus on rights goes awry is when the holder of the right is not clear, and a judge must decide in whom to vest the right (which then leads to a solution). Although the implication of the Coase Theorem is that this choice does not matter, because the efficient solution will obtain regardless, the standard guidance from scholars such as Posner (1973) is for the judge to “mimic the market” and grant the right to the party who values it more highly (and would presumably buy it if it were granted to the other party instead). However, this collapses to a utilitarian basis for assigning rights; even if the right would end up in the hands of the party who valued it more, it should be vested originally in the party who has the best claim to it on moral grounds (such as the basis of property). The dependence of rights on economic valuations rather than a substantive and qualitative claim diminishes their value as rights and once again reduces economics to utilitarianism.

The deeper problem with externalities from a Kantian point of view is that the economic analysis focuses on the *harm* imposed rather than the *wrong* done. Economics, based on brute utilitarianism, treats all harms the same and recommends any measures to make harms efficient. But not all harms are wrongful, and in fact some harms are protected by rights. When a friend and I walk into a crowded coffee shop to have a conversation about ethics and economics, we impose harms on those around us: we take seats that others may have wanted, we may take the last blueberry scones, and most important we add to the din, making it harder for everyone to hear their companions. But no one would claim that we have done anything wrong in the sense of violating anyone's rights. Social interaction invariably leads to incidental harms that do not merit policy attention, and furthermore, any policy implemented to reduce them would violate the rights of the parties involved. A more serious example would be the rights of free speech and assembly valued in most liberal democracies, which protect these activities even if some find them objectionable and harmful in a real sense (and therefore lead to a decrease in total welfare). Such protections in the Bill of Rights to the United States Constitution (and their analogues in other countries) specifically shield rightful behaviors from interference from the state, even if they are lowering welfare.

Only externalities that include violations of rights—those that can be called *wrongful*—deserve attention from welfare economists and policymakers. Coase's classic story of the railroad throwing harmful sparks on neighboring farmers' crops is a perfect example, which he treated as a case of *reciprocal causation*:

The question is commonly thought of as one in which A inflicts harm on B and what has to be decided is: how should we restrain A? But this is wrong. We are dealing with a problem of a reciprocal nature. *To avoid the harm to B would be to inflict harm on A.* (1960: 2, emphasis added)

He does this presumably to focus on the irrelevance of which party holds the right to the efficiency of the eventual solution, but it is clear from his description of the situation that the trains are wrongfully harming the farmers' interests in their crops, and as such the situation merits attention (Epstein 1973). Ironically, though, the element of wrongfulness that raises such situations to this level of importance also suggests that the solution lies not in economics or policy but the law—specifically, tort law, which deals with assigning responsibility for wrongful harms (White 2015c). Tort law addresses the harm as well as the wrongfulness of such acts and focuses on making them “right” rather than efficient. In this sense, tort law respects the rights and interests of harmed parties rather than using them as a means to maximize total welfare by alleviating their harm only up to the point where it stops being worthwhile to do so.

This is not to say that tort law is capable of handling all externalities involving rights violations. The best example is environmental pollution, which violates what Geistfeld calls an “underlying entitlement to physical security” (2014, 389). Economists know all too well the problems with “missing markets” in common resources, while legal scholars focus on the problems with identifying injurers and demonstrating proximate cause, all of which make it very difficult for victims of pollution to seek damages in court. In such cases, economic policy measures such as Pigouvian taxes and direct regulation are indicated. The point is not that tort law always works better than economic solutions, but that only those externalities that involve rights violations deserve officials' attention, which can often be channeled through the existing institution of tort law rather than discretionary and often distortionary economic solutions.

3.4.3 The Nature and Regulation of the Market

The contrast between rights and welfare also extends into the realm of markets and commerce, especially regarding how business activity is understood and regulated by the state. Like all economic activity, firms, industries, and entire markets are judged by economists according to the welfare or total utility generated, and are found lacking to the extent they fall short of the idealized maximum of perfect competition. In such cases, the state intervenes, implementing policy, regulation, or law in an attempt to make the process more competitive and welfare closer to the maximum. Understood this way, all market activity is a mechanism to generate welfare, and any shortcoming in this mechanism—“market failure”—needs to be corrected.²⁸ This is an inherently utilitarian frame which recognizes no rights on the part of market participants to pursue their own

²⁸ On the concept of market failure, see Zerbe and McCurdy (1999).

goals separate from aggregate welfare, while at the same time granting the state an absolute right to manipulate the parameters of market activity to achieve maximal welfare, in which market participants are used simply as means to achieve the ends of the state.

If we focus on rights and autonomy of market participants rather than the consequences of their activity, we can understand the market as a mechanism that coordinates the actions that individual participants—consumers and producers, employees and employers, lenders and borrowers—take in pursuit of their own interests (defined broadly as whatever motivates their choices).²⁹ It allows individuals to make autonomous choices in an environment that constrains their choices only by the equally legitimate choices of other individuals, using prices and other market parameters generated by the aggregated actions of all. In this sense, the operation of the market is judged not by how much total utility it generates but to what degree it allows its participants to pursue their own interests in a mutually consistent manner.³⁰ Market failure would be defined not by suboptimal welfare but by arbitrary limitations or burdens placed on the individual pursuit of interests, interference arising from illegitimate coercion on the part of market participants or external forces such as government actors.

This conception of the market also recognizes that individual market participants have no obligation to contribute to total welfare and have the right to pursue their interests, provided they do so in a way consistent with all others doing the same, regardless of the effect on welfare. This does not rule out laws and moral norms that protect market participants from fraud or deceit, nor does it prevent the state from imposing regulations that protect employees from unsafe workplaces or consumers from dangerous products. These interventions are ultimately grounded in considerations of rights and dignity, not simple utility as such, and as such are a perfectly legitimate concern for the state (which nonetheless must be justified on a case-by-case basis). What would not be legitimate in this Kantian frame are market interventions on the part of the state with the sole intent of increasing total welfare, which transform market participants into mere means to that end.

One of the most vivid examples of this is antitrust law, which aims to increase industrial competition and thereby welfare (White 2016b). To this end, the law penalizes firms for taking actions that counteract these goals of the state, which implies that firms have a legal responsibility to promote total welfare, likening them to agents of the state rather than instruments of citizens pursuing their own interests. Furthermore, antitrust law imposes criminal sanctions, which are normally reserved for actions which not only produce harm but do so wrongfully and to a degree that threatens public safety. Most conceptions of property rights, understood as “bundles” of specific rights, include the right of disposal, which individuals may use to exchange their property for money or other property, including shares of other firms. Some actions, such as restrictions on

²⁹ For more on this conception of interests, see White (2017).

³⁰ The concept of the market as a coordinating mechanism draws from Austrian economics, especially Israel Kirzner’s emphasis on process as opposed to equilibrium (1973); for historical context, see Sautet (2015) as well as, generally, the chapter by Boettke and Woltz in this volume.

terms (such as price) or mergers with other firms, are taken to promote the property owners' interests with predictable negative effects on other parties, such as consumers or competing firms. Consumers have no right to a certain level of prices, nor do other firms have a right to a certain competitive environment, so firms engaging in "anticompetitive" actions are violating no rights that would imply wrongfulness on their part. Unless these actions are wrongful, they are legitimate exercises of underlying and protected property rights, and should not be forbidden by the state simply because they are understood to lower total welfare. In this sense, any harm that arises from them would be akin to nonwrongful externalities, as discussed above: negative impact on third parties created by legitimate activity protected by rights.

In a general Kantian sense, government intervention and regulation involving market activity should be restrained to addressing rights violations and wrongful conduct, such as instances of fraud, deceit, and wrongful injury, rather than attempting to engineer maximal levels of total welfare. To do the latter threatens to trample individual rights and the legitimate exercise thereof. States should not enforce a utilitarian agenda against its citizens, implying an obligation to conform their behavior to the needs of the whole, but only ensure that their autonomously chosen behavior does not wrongfully affect others. The importance and value of the market should be judged by its process rather than its outcome: the latter may be positive, but this is secondary to its function as a mechanism to coordinate the pursuit of interests by individuals constrained by the rights of others.

3.5 CONCLUSION

Kantian deontological ethics provides a counterbalance to the prevalent utilitarianism underlying much of economic practice and its neglect of concepts such as autonomy, dignity, and rights. On the one hand, most economists expressly deny or diminish the importance of moral principles and rights, reducing them to commodified utilitarian considerations, such as property rights that can be dismissed if they fail to maximize welfare. On the other hand, they fail to acknowledge the implicit presence and invaluable contribution of principles and rights, such as when market behavior is assumed to conform to laws and norms. Economists rely on deontological concepts at the same time they ignore their importance and contribution to theory, practice, and policy.

I do not argue that Kantian ethics or deontology must replace utilitarianism in the economics toolbox. Both are necessary and valuable, both in terms of descriptive theory as well as prescriptive policy analysis. Economists cannot model, understand, or predict individual behavior accurately without taking into account the ways people include moral principles and duties into their decision-making. By the same token, economists recommending policies, regulations, and laws designed to increase well-being must not do so at the expense of human rights or dignity. Through its grounding in autonomy and dignity, Kantian ethics reminds us of the value of the person, a concept that is all too

easily lost when economists try to quantify, formalize, and aggregate human experience. This is the utilitarian legacy of mainstream economics, which I argue needs to be supplemented with a strong emphasis on the worth of the individual, which is perhaps the greatest lesson to be learned from Immanuel Kant.

REFERENCES

- Akerlof, George A. 1991. "Procrastination and Obedience." *American Economic Review Papers and Proceedings* 81: 1–19.
- Andreoni James. 1990. "Impure Altruism and Donations to Public Goods: A Theory of Warm-Glow Giving." *Economic Journal* 100: 464–77.
- Backhouse, Roger E., and Steven G. Medema. 2009. "Defining Economics: The Long Road to Acceptance of the Robbins' Definition." *Economica* 76: 805–20.
- Baumeister, Roy F., and Todd F. Heatherton. 1996. "Self-Regulation Failure: An Overview." *Psychological Inquiry* 7: 1–15.
- Baumol, William J. 1972. "On Taxation and the Control of Externalities." *American Economic Review* 62: 307–22.
- Becker, Gary S. 1976. *The Economic Approach to Human Behavior*. Chicago, IL: University of Chicago Press.
- Becker, Gary S. 1981. *A Treatise on the Family*. Cambridge, MA: Harvard University Press.
- Bergstrom, Theodore C. 1996. "Economics in a Family Way." *Journal of Economic Literature* 34: 1903–34.
- Betzler, Monika (ed.). 2008. *Kant's Ethics of Virtue*. Berlin: Walter de Gruyter.
- Buchanan, James M. 1990. "The Domain of Constitutional Economics." *Constitutional Political Economy* 1: 1–18.
- Cathcart, Thomas. 2013. *The Trolley Problem—or—Would You Throw the Fat Guy Off the Bridge?* New York: Workman.
- Coase, Ronald H. 1960. "The Problem of Social Cost." *Journal of Law and Economics* 3: 1–44.
- Coleman, Jules L. 1984. "The Foundations of Constitutional Economics." In Richard B. MacKenzie (ed.), *Constitutional Economics: Containing the Economic Powers of Government* (Lexington, MA: Lexington Books), pp. 141–55.
- Coleman, Jules L. 1985. "Crimes, Kickers and Transaction Structures." In J. Roland Pennock and John W. Chapman (eds.), *Criminal Justice: NOMOS XXVII* (New York: New York University Press), pp. 313–28.
- Cureton, Adam, and Thomas E. Hill. 2015. "Kant on Virtue and the Virtues." In Nancy E. Snow (ed.), *Cultivating Virtue: Perspectives from Philosophy, Theology, and Psychology* (Oxford: Oxford University Press), pp. 87–109.
- Dworkin, Ronald M. 1977. *Taking Rights Seriously*. Cambridge, MA: Harvard University Press.
- Dworkin, Ronald M. 1980. "Is Wealth a Value?" *Journal of Legal Studies* 9: 191–226.
- Edmonds, David. 2014. *Would You Kill the Fat Man? The Trolley Problem and What Your Answer Tells Us about Right and Wrong*. Princeton, NJ: Princeton University Press.
- Engstrom, Stephen, and Jennifer Whiting (eds.). 1996. *Aristotle, Kant, and the Stoics: Rethinking Happiness and Duty*. Cambridge: Cambridge University Press.
- Epstein, Richard A. 1973. "A Theory of Strict Liability." *Journal of Legal Studies* 2: 151–204.
- Etzioni, Amitai. 1987. "Toward a Kantian Socio-Economics." In Charles K. Wilber (ed.), *Economics, Ethics, and Public Policy* (Lanham, MD: Rowman & Littlefield), pp. 139–49.

- Etzioni, Amitai. 1988. *The Moral Dimension: Toward a New Economics*. New York: Free Press, 1988.
- Fine, Ben. 2002. "‘Economic Imperialism’: A View from the Periphery." *Review of Radical Political Economics* 34: 187–201.
- Foot, Philippa. 1967. "The Problem of Abortion and the Doctrine of Double Effect." *Oxford Review* 5: 5–15.
- Gaus, Gerald F. 2001a. "What Is Deontology? Part One: Orthodox Views." *Journal of Value Inquiry* 35: 27–42.
- Gaus, Gerald F. 2001b. "What Is Deontology? Part Two: Reasons to Act." *Journal of Value Inquiry* 35: 179–93.
- Geistfeld, Mark A. 2014. "The Tort Entitlement to Physical Security as the Distributive Basis for Environmental, Health, and Safety Regulations." *Theoretical Inquiries in Law* 15: 387–415.
- George, David. 2001. *Preference Pollution: How Markets Create the Desires We Dislike*. Ann Arbor, MI: University of Michigan Press.
- Gregor, Mary J. 1963. *Laws of Freedom: A Study of Kant’s Method of Applying the Categorical Imperative in the Metaphysik der Sitten*. Oxford: Basil Blackwell.
- Grüne-Yanoff, Till, and Sven Ove Hansson (eds.). 2009. *Preference Change: Approaches from Philosophy, Economics and Psychology*. Dordrecht: Springer.
- Herman, Barbara. 1989. "Murder and Mayhem: Violence and Kantian Casuistry." *The Monist* 72: 411–32.
- Hicks, John R. 1939. *Value and Capital*. Oxford: Clarendon Press.
- Hill, Thomas E., Jr. 1989. "The Kantian Conception of Autonomy." In John Christman (ed.), *The Inner Citadel: Essays on Individual Autonomy* (Oxford: Oxford University Press), pp. 99–105.
- Hill, Thomas E., Jr. 1991. *Autonomy and Self-Respect*. Cambridge: Cambridge University Press.
- Hiller, Janine S. 2013. "The Benefit Corporation and Corporate Social Responsibility." *Journal of Business Ethics* 118: 287–301.
- Holmes, Oliver Wendell, Jr. 1897. "The Path of the Law." *Harvard Law Review* 10: 457–78.
- Irwin, Terence. 2004. "Kantian Autonomy." In John Hyman and Helen Steward (eds.), *Agency and Action* (Cambridge: Cambridge University Press), pp. 137–64.
- Jost, Lawrence, and Julian Wuerth (eds.). 2011. *Perfecting Virtue: New Essays on Kantian Ethics and Virtue Ethics*. Cambridge: Cambridge University Press.
- Kamm, F. M. 2016. *The Trolley Problem Mysteries*. Oxford: Oxford University Press.
- Kant, Immanuel. 1785. *Grounding for the Metaphysics of Morals*. Trans. James W. Ellington. Indianapolis, IN: Hackett (1993 edition).
- Kant, Immanuel. 1781/1787. *Critique of Pure Reason*. Trans. Norman Kemp Smith. New York: St. Martin’s Press (1929 edition).
- Kant, Immanuel. 1788. *Critique of Practical Reason*. Trans. Lewis White Beck. Upper Saddle River, NJ: Prentice Hall (1993 edition).
- Kant, Immanuel. 1797. *The Metaphysics of Morals*. Trans. and ed. Mary J. Gregor. Cambridge: Cambridge University Press (1996 edition).
- Kant, Immanuel. 1799. "On a Supposed Right to Lie Because of Philanthropic Concerns." In *Grounding for the Metaphysics of Morals*, trans. James W. Ellington (Indianapolis: Hackett Publishing, 1993 edition), pp. 63–7.
- Kirzner, Israel M. 1973. *Competition & Entrepreneurship*. Chicago, IL: University of Chicago Press.
- Kronman, Anthony T. 1980. "Wealth Maximization as a Normative Principle." *Journal of Legal Studies* 9: 227–42.

- Margolis, Howard. 1984. *Selfishness, Altruism, and Rationality*. Chicago, IL: University of Chicago Press.
- Mercuro, Nicholas, and Steven G. Medema. 2006. *Economics and the Law: From Posner to Post-Modernism and Beyond*. 2nd ed. Princeton, NJ: Princeton University Press.
- Minkler, Lanse. 2008. *Integrity and Agreement: Economics When Principles Also Matter*. Ann Arbor, MI: University of Michigan Press.
- Moore, Michael S. 1989. "Torture and the Balance of Evils." *Israel Law Review* 23: 280–344.
- O'Donoghue, Ted, and Matthew Rabin. 2001. "Choice and Procrastination." *Quarterly Journal of Economics* 116: 121–60.
- O'Neill, Onora. 2004. "Kant: Rationality as Practical Reason." In Alfred R. Mele and Piers Rawling (eds.), *The Oxford Handbook of Rationality*, (Oxford: Oxford University Press), pp. 93–109.
- Paton, H. J. 1947. *The Categorical Imperative: A Study in Kant's Moral Philosophy*. Philadelphia, PA: University of Pennsylvania Press.
- Pigou, A. C. 1932. *The Economics of Welfare*. 4th ed. London: Macmillan.
- Posner, Richard A. 1973. *Economic Analysis of Law*. Boston, MA: Little, Brown and Company.
- Posner, Richard A. 1983. *The Economics of Justice*, 2nd ed. Cambridge, MA: Harvard University Press.
- Rawls, John. 1971. *A Theory of Justice*. Cambridge, MA: Harvard University Press.
- Robbins, Lionel. 1932. *An Essay on the Nature and Significance of Economic Science*. London: Macmillan and Co.
- Ross, W. D. 1930. *The Right and the Good*. Oxford: Oxford University Press.
- Sautet, Frederic. 2015. "Market Theory and the Price System." In Peter J. Boettke and Christopher J. Coyne (eds.), *The Oxford Handbook of Austrian Economics* (Oxford: Oxford University Press), pp. 65–93.
- Scheffler, Samuel (ed.). 1988. *Consequentialism and Its Critics*. Oxford: Oxford University Press.
- Sen, Amarta K. 1977. "Rational Fools: Fools: A Critique of the Behavioral Foundations of Economic Theory." *Philosophy & Public Affairs* 6: 317–44.
- Sherman, Nancy. 1997. *Making a Necessity of Virtue: Aristotle and Kant on Virtue*. Cambridge: Cambridge University Press.
- Sullivan, Roger J. 1989. *Immanuel Kant's Moral Theory*. Cambridge: Cambridge University Press.
- Sullivan, Roger J. 1994. *An Introduction to Kant's Ethics*. Cambridge: Cambridge University Press.
- Thomson, Judith Jarvis. 1976. "Killing, Letting Die, and the Trolley Problem." *The Monist* 59: 204–17.
- White, Mark D. 2011a. *Kantian Ethics and Economics: Autonomy, Dignity, and Character*. Stanford, CA: Stanford University Press.
- White, Mark D. 2011b. "Pro Tanto Retributivism: Judgment and the Balance of Principles in Criminal Justice." In Mark D. White (ed.), *Retributivism: Essays on Theory and Policy* (Oxford: Oxford University Press), pp. 129–45.
- White, Mark D. 2015a. "Judgment: Balancing Principle and Policy." *Review of Social Economy* 73: 223–41.
- White, Mark D. 2015b. "A Kantian-Economic Approach to Altruism in the Household." *Palgrave Communications* 1, Article no. 15005: 1–6.
- White, Mark D. 2015c. "On the Relevance of Wrongfulness to the Concept of Externalities." *Æconomia* 5: 313–29.

- White, Mark D. 2016a. "The Virtues of a Kantian Economics." In Jennifer A. Baker and Mark D. White (eds.), *Economics and the Virtues: Building a New Moral Foundation* (Oxford: Oxford University Press), pp. 94–115.
- White, Mark D. 2016b. "On the Justification of Antitrust: A Matter of Rights and Wrongs." *Antitrust Bulletin* 61: 323–35.
- White, Mark D. 2017. "Preferences All the Way Down: Questioning the Neoclassical Foundations of Behavioral Economics and Libertarian Paternalism." *Economia* 7: 353–73.
- Zerbe, Richard O., Jr., and Howard E. McCurdy. 1999. "The Failure of Market Failure." *Journal of Policy Analysis and Management* 18: 558–78.

CHAPTER 4

ETHICAL PLURALISM IN ECONOMICS

JONATHAN B. WIGHT

4.1 INTRODUCTION

PLURALISM in economics is an approach that draws on more than one paradigm for understanding the functioning of families, markets, and the public sphere, and for normatively assessing the effectiveness of institutions and policies using a variety of methodologies and values. Pluralism in the ethical dimensions of economics draws on the narrower discussion of how people use competing and complementary moral frameworks to understand individual action and to evaluate public policies.

A positive neoclassical analysis of families, for example, might model household members as utility maximizers with pre-formed and stable preferences, who make rational decisions to achieve personal outcomes. A positive pluralist approach might highlight the conflicted nature of human experience, the complexity of preference formation in a social context, and the potential conflicts between identity, commitment, and outcomes. A promise or duty, not a rational calculation of consequences, may lead some people to behave in ways that are other regarding in certain settings: this could explain tipping at highway restaurants or devoting time and resources for the well-being of others. Caregivers often make sacrifices for family members and friends (Folbre 2012); some of the recipients may carry on a genetic line of the altruist but in many cases do not. Or, people may feel compelled to take actions based on principles about what is right and virtuous, not from an anticipation of particular outcomes. A pluralist might thus explore how behaviors are at times driven by instinctual or learned social behaviors and norms, inspired by moral sentiments, and not always by the purely rational calculations of self-interest. Without a pluralist conception of ethics, economists are hard pressed to argue for ethics at all—even when describing their own professional conduct.

This chapter provides an analysis of how the neoclassical economic approach to human behavior and to human welfare, concerned with outcome goals and evaluative methods, is enhanced by drawing upon concepts arising out of duty ethics and virtue ethics. In this light, the moral frameworks are not always in opposition, but often complement each other. As elaborated in Section 4.3, the neoclassical welfare approach to maximizing the economic surplus through trade turns out to rely on supporting institutions for upholding principles of human rights, and the cultivation of individuals who are called on to uphold these principles. Economic efficiency thus builds on a larger world of ethical pluralism (Wight 2017). The main approaches to understanding morality—evaluating outcomes, committing to rules and to principles of duty, and adhering to notions of virtue—constitute three ways of framing conduct in a social setting. Ethical pluralism is the claim that all three approaches are to some degree useful to positive economics because each provides insights into human behavior. Each is also useful in normative economics, because a narrow approach has limitations that can be solved only by introducing elements of other frameworks.

4.2 BRIDGING ETHICAL BOUNDARIES

A pluralist bridges ethical borders both to understand human behavior and to make normative claims about it.

4.2.1 Horizontal and Vertical Pluralism

Horizontal pluralism crosses the boundaries of the economic actor, the action itself, and the resulting outcomes of that act, as shown in Figure 4.1.

Both a positive and a normative analysis could proceed from any of these starting points. An economic *actor* can have an identity and can gain meaning from a life adhering to ideals of virtue within a particular social setting (focusing on the actor in the leftmost column). At the same time, every social context introduces formal and informal norms and rules of behavior that constrain the choice of *action* (focusing attention on the middle column). The rightmost column identifies possible *outcomes* that can be chosen as the desired goal (the traditional neoclassical economic focus). A horizontal pluralist might thus claim that the motives of the actor, as well as the constraints on actions arising from her duties or commitments, offer additional insights into economic choice problems that an outcome-based approach alone cannot. The claim of horizontal pluralism is that each moral framework has limitations that can be overcome by introducing concepts that bridge across the relevant approaches.

Vertical pluralism identifies different traditions within each type of moral framework. Outcome-based ethical choice could focus on maximizing pleasure and minimizing pain (utilitarianism), maximizing the economic surplus (efficiency), maximizing fairness,

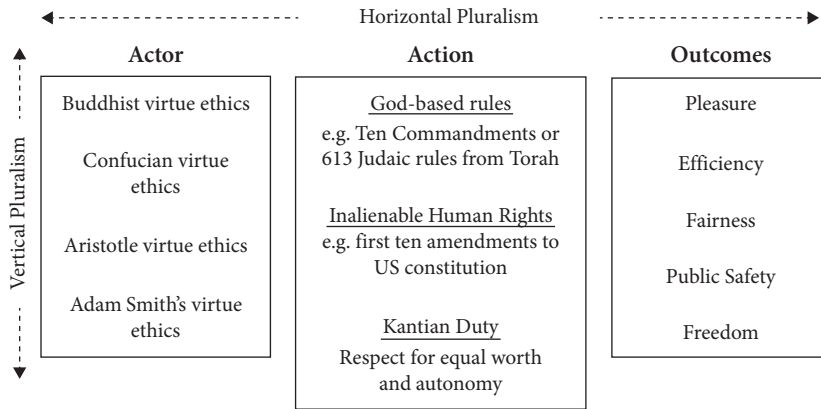


FIGURE 4.1 Multiple approaches to understanding right from wrong.

freedom, public safety, or some other value, or achieving some combination of these. Most people often care about more than one type of outcome, and a vertical pluralist might be willing to trade off some efficiency to get more fairness. A “satisficing” outcome-based pluralist would want certain minimum amounts of some of these results, which could be achieved only by abandoning the focus on maximization of a single outcome.

Vertical pluralism within the middle column of duty and rules considers the notion that the action itself is the focus of analysis, rather than the outcomes produced by it. These might appeal to God-based commandments, human rights, or to Kantian categorical imperatives. Rather than obeying commandments given from God, Kant asks us to use the rational minds that God provides us with to arrive at a principle of acting. That principle is not to make an exception of ourselves and to be able to will into existence a world in which our actions could be universalized.

Vertical pluralism is also evident in the oldest moral framework, virtue ethics, which deals with the motives and character of the actor. This method of analysis dates back thousands of years in oral and written traditions. In India, Siddhartha Gautama sought to purify the human character by minimizing wants. In modern parlance, Buddhist philosophy teaches us to modify our preferences to have no preferences, and the economic actor develops a character of self-control, preferring less to more consumption. In China, Confucius developed a hierarchical community structure that builds on virtuous concepts of benevolence, propriety, and justice up and down the social ladder. In ancient Greece, Aristotle encouraged excellence of character that would result in *eudaimonia*, or human flourishing. Virtue from this perspective is necessarily pluralist, because virtues cannot be traded off against one another: prudence, temperance, and justice must each and separately be properly balanced between the vices of deficiency or excess. This harmonizing requires practical reasoning that is not the same thing as maximizing particular outcomes.

John Stuart Mill refines Bentham’s monist view of utilitarian goodness (deriving from a single measure of net “pleasure”) to a pluralist view: “some kinds of pleasure

are more desirable and more valuable than others” (1863: 8). Pleasures arising from the “higher” faculties, which include mental, spiritual, aesthetic, and moral aspects, should be treated differently, which reflects an acceptance of vertical pluralism and multiple output dimensions. But to differentiate higher from lower pleasures requires a person of experience and virtue, which is an acceptance of some degree of horizontal pluralism across moral frameworks. Mill admits that “Utilitarianism, therefore, could only attain its end by the general cultivation of nobleness of character . . .” (ibid.: 11). The most influential economist of the mid nineteenth century had a thoroughly pluralist outlook.

4.2.2 Modeling Markets and Human Behavior

Many modern neoclassical economists implicitly assume that human behavior is more ethically varied than commonly supposed. The economic agent’s standard utility function posits maximization of utility for self, which may or may not include other regarding or altruistic preferences. Unpacking how this might work in practice requires the adoption of complementary ethical frameworks. For example, the existence and operation of firms is often ignored in microeconomic models of resource allocation. It is often simply assumed that businesses maximize profits, and in so doing, allocate resources to produce products that best satisfy consumer preferences at the lowest possible cost (making all the usual assumptions). Introducing the business firm—and its manager-agents who will carry out this coordination activity—can lead to distressing results because of well-known principal-agent problems. Thus, when Milton Friedman writes that, “Few trends could so thoroughly undermine the very foundations of our free society as the acceptance by corporate officials of a social responsibility other than to make as much money for their stockholders as possible” (1962: 133), most people interpret this as an attack on stakeholder theory (for example, workers and the environment). A different interpretation is that this is also a recognition of pluralist ethics. The reasoning is as follows: managers, Friedman claims, are expected to make as much profit as possible for owners, but doing so within the “basic rules of the society, both those embodied in law and those embodied in ethical custom” (1970). This statement is a wide vehicle for introducing pluralist ethics into the firm’s production function.

“Ethical custom” is a fluid concept that requires elaboration through dialogue to understand and implement, and it evolves over time. In Friedman’s context, ethical custom would include the fiduciary duty of a manager to safeguard shareholder interests and to disregard his or her own competing private interests. Persons of character uphold their duty to others not simply because of enlightened self-interest (hoping to get repeat business from the appearance of fair play) but because of having internalized the notion that fair play is the right action, regardless of the consequences. In Adam Smith’s view, from *The Theory of Moral Sentiments* (1759), a person of virtue desires to be *worthy* of admiration; virtue is not an instrumental response to incentives.

Friedman's formative early years likely led him to interact with business people who professed high ideals along these lines and who showed self-restraint when needed (both of his parents were merchants). In the early to mid twentieth century business leaders were participants in a number of philanthropic service organizations and clubs such as the Kiwanis and the Rotary International (both established in Chicago in 1905); the Lions (established in Detroit in 1915); the Jaycees (the Junior Chamber of Commerce, established in St. Louis in 1920); and the Better Business Bureau (established in Atlanta in 1912). These groups all promote virtue ethics, fair dealing, and service to others. Friedman, in endorsing capitalism, likely thinks this version of capitalism prevails, namely that many or most business leaders have sufficient virtue ethics to put shareholder interests ahead of personal interests, and that the instrumental value of a market system and profits can be salvaged despite some bad apples who abuse the system for their own gain. In short, Friedman must believe that the norms of duty (embedded in ethical custom) are strong enough to withstand the forces of selfish individualism. If humans behave according to these pluralist beliefs or mindsets, then economic agents are at times guided in their choices by considerations about the nature of virtue, and the commands of duty, in addition to calculating personal outcomes.

Amartya Sen, another Nobel Laureate, embraces pluralist ethics quite explicitly. With regard to horizontal pluralism, Sen draws our attention to the deontological perspective by arguing that people's commitments or promises, such as to truth telling, are internal moral constraints on maximizing behavior. Thus, "choice may reflect a compromise among a variety of considerations of which personal welfare may be just one" (1977: 324). While it is possible to insert duty and virtue into a utility function, it mixes apples and oranges, and different ethical frameworks are generally not reducible to a consequentialist framework. Adam Smith notes that such an approach is "contrived" and produces ethical behavior of a "much inferior order" (1759: 263). Second, in regard to vertical pluralism, Sen addresses multiple outcome goals in normative economics, a subject covered in the following section.

Vernon Smith, who won the Nobel Prize for his work in experimental economics, likewise has embraced a pluralist ethic for understanding some human behavior. In particular, he finds that instinctive and affective moral sentiments operate to generate trust and promote cooperation in market settings, as theorized by Adam Smith (McCabe et al. 2003; Smith 2013). Paul Zak (2009) further identifies a physiological basis for moral sentiments in social settings that is not tied to conscious or rational maximizing behaviors.

A number of researchers have drawn larger narratives from these and other findings, including Sen (1987), Frey and Oberholzer-Gee (1997), Gintis et al. (2005), Bowles and Gintis (2011), Gneezy et al. (2011), and Bowles (2016). Moral constraints on market activity are modeled to arise for reasons of civic virtue and moral sentiments, in addition to enlightened self-interest (Wight 2015). Exposing students only to one-sided treatments of *homo economicus* may weaken the natural expectation of social reciprocity and trust in some settings (Frank et al. 1993; Nelson 2016).

4.3 CONSEQUENTIALIST EVALUATIONS OF HUMAN WELFARE

Consequentialism examines the outcomes of an action or policy to judge its normative worth or lack thereof. There are varieties of consequentialism, of which the neoclassical approach is a hybrid.

4.3.1 Neoclassical Welfare Theory

Neoclassical economists take utilitarianism as an historical precursor, but differ so much from this approach today that calling it utilitarian seems a misnomer. Neoclassical welfare economists reject the notion that Bentham's pleasure and pain, or Mill's higher and lower states of mind, could be useful concepts for evaluating an economic system. Striving for greater objectivity, they adopted in the twentieth century an approach to human welfare built on satisfying preferences, which essentially has no substantive content. Consumers "reveal" their preferences in the market, and economists do not know or care what those preferences actually are. Josiah exchanges \$10 for a pound of steak: perhaps he loves meat and will get great pleasure from it, but pleasure is not the metric of analysis. Josiah could equally be a masochist who inflicts pain on himself or a sadist who inflicts pain on vegetarians. Neoclassical welfare theorists have no concern about the actual mental states of mind involved, nor for the substantive consequences of eating a steak in terms of nutrition and health; the dollar votes cast in the marketplace are the only metric that matters.

In neoclassical welfare theory the goal is to satisfy the most preferences possible, measured through the proxy of maximizing the economic surplus through trade. While economists sometimes slip into speaking about pleasure or happiness as the result of satisfying preferences, no such interpretation is permitted, because it leads to the heresy of interpersonal comparisons of utility. Neoclassical economists are decidedly non-utilitarian in how they answer the question of "what" outcome or consequence to value.

Neoclassical economists also differ markedly from utilitarian views in deciding whose interests count in a normative analysis. Someone who is indigent and has no buying power in the market is said to have no interests that should be considered in assessing the efficiency of the economic system. Indeed, the first fundamental welfare theorem in neoclassical economics states that under ideal institutional circumstances, a free market would create an efficient allocation of resources without caring at all about distribution. This led Sen to complain that:

An economy can be optimal in this sense even when some people are rolling in luxury and others are near starvation as long as the starvers cannot be made better off without cutting into the pleasures of the rich.... In short, a society or an economy can be Pareto-optimal and still be perfectly disgusting. (1970: 22)

This criticism is tempered by the second fundamental welfare theorem, which states that lump-sum redistributions of endowments to indigents would produce a different combination of market clearing prices and outputs, and efficiency would still be maintained even though prices and outputs could change substantially with redistributions. The acknowledgment that the economy can still be microeconomically efficient even after redistribution is not the same thing as arguing that economic justice is an important outcome that should be pursued. Treating the interests of people differently depending on their spending power creates a philosophical conflict with a democratic system in which every citizen is presumed to be treated equally under the law. Resolving this conflict requires, at minimum, the use of vertical and horizontal pluralism by incorporating considerations of fairness, due process, and voice when discussing public policies that inflict harm on some people (a point developed below).

With regard to aggregation, neoclassical economists sum results across individuals by counting all the dollars votes cast in the marketplace, treating each person's dollar as equal to any other person's. It doesn't matter if the consumer is a maid or a matriarch, a citizen or an illegal immigrant, equal status is accorded in the welfare economist's bazaar. Bertolt Brecht, who had no love of capitalism, understood this point: "In business you ask what price, not what religion. And Protestant trousers keep you just as warm" (1941: 52–3). By saying \$100 spent by a wealthy person adds equivalently to human welfare as \$100 consumed by a pauper, economists deny the utilitarian view that the substantive outcomes of spending money involve diminishing marginal benefits, and can be (roughly) compared between individuals for the purpose of enacting public policy. However, no such states of mind can be inferred from the neoclassical position, which avoids making any interpersonal comparisons of mental states.

4.3.2 Pareto Efficiency

A market outcome is judged to be more or less "efficient" depending on the answer to this question: "Is it possible through voluntary trade to improve the satisfaction of one person's preferences without hurting anyone else's preference satisfaction?" This standard, expounded by Vilfredo Pareto in 1906, is widely used in textbooks and offers insights into resource problems implicitly developed by Edgeworth in 1881. From a point of maximum *ophelimity* or maximum economic utility, "it is not possible to move away from [it] helping, or harming, both individuals at one and the same time; but necessarily, if it is agreeable to the one, it is disagreeable to the other" (Pareto 1906: 262). Yet why should voluntary trade be upheld as "the" moral standard upon which to judge the efficacy of the economic system? Virtually no textbooks address this issue, and it often remains an implicit ethical underpinning. More explicitly, one could argue that voluntary trade is not justified because of any specific outcomes produced, since economists say they are not interested in specific outcomes or in any actual mental states that might emerge. Rather, the Pareto approach can be justified because it recognizes each person's right to autonomy, to self-determination through rational choice. People

should have the freedom to make choices or trades according to their own reasons and motives, without manipulation or coercion. This justification is thoroughly Kantian, relying on the acceptance of others' rights and the corollary acceptance of our duties to respect others in a certain way. The justification for the Pareto test for efficiency is not based on a consequentialist ethic; it relies on Kantian duty ethics (or something similar in libertarian ethics) for its formulation. This is an example of horizontal pluralism within the very foundations of neoclassical welfare theory.

The Pareto test is a useful construct for solving some practical resource problems, such as in allocating seats on a full airplane. Non-market solutions, such as first-come, first-served, rationing, lottery, and favoritism typically fail the Pareto test, because individuals who value the air flight more than cash are unable to reliably swap for it, and those who value a cash reward rather than the flight are likewise unable to reliably satisfy that preference. Allowing an airline to solicit volunteers to trade travel time today for a reward later usually works without conflict, according respect to passengers in a process that treats each as autonomous persons of dignity; along the way, it also optimizes preference satisfaction.

A recent case, however, demonstrates why Pareto's test cannot stand alone as an ethical approach to allocation. A United Airlines passenger, Dr. David Dao, was dragged from his seat after he declined the offer of a voluntary swap of his seat on a flight from Chicago to Louisville. He suffered injuries including a concussion, a broken nose, and loss of two front teeth. The initial problem with the Pareto approach in this case is that it was not pursued far enough—passengers were offered \$1,000 to wait 24 hours—but no one on the plane took the bait. Airline attendants subsequently used force to dislodge Dr. Dao from his seat and the aircraft. After Dr. Dao's settlement with the airline, United agreed to raise its highest potential reward to \$10,000 for passengers who voluntarily switch travel times. The CEO of United Airlines, who presumably wants to make profits, stated, "Every customer deserves to be treated with the... deepest sense of dignity and respect" (BBC News 2017). Although the employees who evicted Dr. Dao were focusing on actions that would achieve profit-outcome goals, the CEO's acknowledgement draws on a completely different aspect of ethics, namely duty and virtue.

4.3.3 Kaldor-Hicks Efficiency

The larger problem identified in this situation is that Pareto's solution to allocation is woefully inadequate for solving most policy problems, for two reasons. First, payments to losers to get their voluntary cooperation for a public policy change are generally not possible because of the high costs of assessing winners and losers and transacting payments. A late 1930s revision of welfare theory tries to overcome this problem: it asserts that as long as the winners gain more than the losers lose, the policy is efficient at maximizing the economic surplus, even when there is no compensation paid. This approach relies on government coercion (threatening to punish economic agents if they do not

conform) or compulsion (using force to compel particular actions) to ensure that the economic surplus is maximized.

This Kaldor-Hicks reformulation throws out the window Pareto's concern for due process and for treating economic agents with respect. The approach is decidedly non-pluralist, which gives rise to numerous problems. If the economic surplus is the only consideration, then a heavy-handed state can impose resource reallocations that enhance efficiency by arresting or killing protestors, suspending human rights, blocking democratic reforms, and prohibiting free speech. The potential for such abuses is what made the *faux* memo at the World Bank about the efficiency gains of moving polluting industries to African dictatorships so "toxic" (Rosenberg 2001). Kaldor-Hicks efficiency that entails some coercion or compulsion is justifiable only when there is a context of protection of human and property rights, the rule of law, an impartial judiciary, democracy, and a free press.

Closer to home, the Supreme Court's *Kelo* case in 2005, in which a woman was forced to sell her home so a private developer could in theory make greater profit, demonstrates the weakness of property rights when jurists buy the Kaldor-Hicks formulation of efficiency. The notion that impartial economic experts can—without reference to a pluralist conception of property rights or procedures arising out of Kantian ethics or human rights—make sweeping resource allocation judgments is terrifying on the face of it. Deferring to economic experts on such matters represents a threat to the democratic way of life, as elaborated in Levy and Peart's *Escape from Democracy: The Role of Experts and the Public in Economic Policy* (2016). In defense of Kaldor-Hicks, most economists would likely assert that their models implicitly assume something like what Frank Knight said a half-century ago: "It is well to state explicitly at the outset that the society considered in this essay is the sovereign democratic state, that is, a modern Western nation, where law is made and enforced by a responsible government within the context of representative institutions" (1960: 19). This is a pluralist ethical context in which rule of law, not outcomes alone, justify the notion of efficiency.

4.3.4 Other Goals and Methodologies

Even if payments to policy losers were possible and accomplished, attributing all human welfare to gains or losses in preference satisfaction is a stretch. Other outcomes beside the economic surplus matter to well-being. Economists often draw on vertical pluralism to highlight those other outcomes. Many economists endorse patent protection, for example, even though the policy will create a temporary monopoly and a deadweight loss of economic surplus in today's market. The rationale for patents is that they will stimulate creativity, risk-taking, and innovation, giving rise to new products and processes in the future. This will, presumably, enhance preference satisfaction in some future time period, and the consumers of that day may not yet be born. Arguing in favor of patents therefore relies on a normative approach in which consumers living in the

future receive a higher weighting of their interests than consumers living in the present: *dynamic* efficiency trumps *static* efficiency. Environmental economists and macroeconomists encounter and address similar issues of intergenerational justice, often using ad hoc methods of vertical pluralism to select goals, values, and methods of analysis.

Similarly, when Becker and Elias (2007) argue for creating a market in human organs, the goal they promote is not economic efficiency, but that of saving lives. Why should saving lives have priority in this case, and not in other cases of allocation? In this case, saving lives and economic efficiency may be positively correlated, but this cannot always be assumed. Health care spending in the United States on cosmetic plastic surgery (such as breast augmentation) grew by 132 percent between 2000 and 2016, reaching \$8 billion per year by 2016 (Gould and Mosher 2017). Such spending satisfies the preferences of consumers who have the financial resources to command those medical services. At the same time, life expectancy declined in certain segments of the population (Kochanek et al. 2017), some of it presumably because of lack of access to timely medical intervention. Achieving economic efficiency in satisfying preferences does not always mean achieving medical efficiency in saving lives—in fact, there are strong a priori reasons for worrying about this claim (Arrow 1963; Reinhardt 2001). Economists often glide between different outcome goals without stopping to contemplate how vertical pluralism is necessary and can be explicitly justified using ethical theory.

As discussed earlier, Amartya Sen argues that the standard way of modeling human welfare through preference satisfaction is highly limited. A woman who has been denied education and opportunities may not have a preference for reading, yet to ignore her lack of capabilities is a serious problem for welfare assessment. Rather than the standard instrumental theory of money providing the only measure of welfare, Sen developed (with Martha Nussbaum) a capabilities approach that identifies substantive measures of well-being (Sen 1999). Building on this, Sen, with economist Mahbub ul Haq, created the Human Development Index (HDI), a widely used hybrid of instrumental and substantive outcome measures. Along similar lines, five Nobel laureates in economics proposed a more nuanced account of how economic activity affects social progress along dimensions of sustainability (Stiglitz et al. 2009).

There are additional values to consider when evaluating an economic system or public policy. These would include freedom, fairness, public safety, and so on. More broadly, some of the positive attributes of markets or policies have nothing to do with outcomes per se, but rather with process and with character. John Stuart Mill speculated that the “economical advantages of commerce are surpassed in importance by those of its effects which are intellectual and moral” (1848: 174).

In addition to pluralist outcomes, pluralist methodologies also find their way into economic evaluation. Arjo Klamer, in evaluating the role of the arts in society, argues that preferences cannot be assumed to be exogenous to the art market and to the social process by which a community of artists, critics, and private and public funders come to see it. He notes:

Art is social. Its values come about in a social setting. A singular work derives its values from being a contribution to a common practice, that is, a practice that a group of

people, artists, connoisseurs, museum directors have in common. That is why art is a shared good. Such a good does not appear in standard economics. (2016: 371)

Similar to a community of artists, economists themselves create a shared good of theory and a common practice of research. The preferences of economists are, to some extent, shaped by their interactions with others, and by a concern for proper action and process.

4.4 ECONOMISTS AS ETHICISTS

When economists discuss their own internal procedures and standards for judging right from wrong, they adopt a pluralist ethical framework.

4.4.1 Code of Professional Conduct

In early 2018 the American Economic Association issued a draft “Code of Professional Conduct,” which upholds principles of intellectual and professional integrity to guide economists in every facet of employment. The drafting of the Code was a reaction to reports of sexual stereotyping and harassment on a job information-sharing website (Wu 2017); pressure had been mounting for years for there to be such an official deliberation (DeMartino 2011). The AEA’s highest purpose is said to be the encouragement of economic research (an outcome-based value). Yet to achieve the best outcomes, the draft Code of Conduct makes an appeal to deontological ethics, insisting that professionals have a duty to act in certain ways toward others. Good outcomes “demand honesty and transparency in conducting and presenting research, disinterested assessment of ideas, and disclosure of conflicts of interest” (AEA 2018). The word “obligation” is used to emphasize the duty to others that should not be abridged to achieve some goal. “This obligation applies even when participating anonymously,” presumably on forums like the “Economics Job Market Rumors” website that sparked the creation of the code.

Yet if economists are supposed to police themselves, professionals must develop the minimum self-control needed to abide by rules that may not align with their own interests. This means that the purveyors of the Code of Conduct (and its critics) have in mind something deeper about the development of character, drawing on the virtue ethics framework. Without self-control, obligations and duties are meaningless. Moreover, the principles identified in the Code are deliberately “parsimonious” (Campbell 2018), ostensibly relying on the virtue of individuals to translate the vague and incomplete language into habits of character.

To summarize, while economists use outcome-based models (often drawing upon selfish individualism) for predicting and evaluating other sectors of the economy, they propose duty-based obligations in economics research; they rely on the self-control of individuals to generate a culture of openness and integrity necessary for producing

high-quality outputs. If this sounds contradictory and convoluted, it should, because the AEA committee's was ad hoc, offering a quick fix for the longstanding omission in the association of explicit attention to duty and virtue in the economic process. Nevertheless, it is a tacit admission that a pluralist approach to ethics in economics is a necessary condition for starting the analysis of a complex social organization.

4.4.2 Why a Code is Not Sufficient

Not all economists are happy with the notion of a code of ethics, perhaps believing in the invisible hand of self-regulation in the market. DeMartino and McCloskey (2018) point out the weakness of a code in the absence of an extended conversation about its meaning and application. While the AEA draft code focuses on creating an open environment for the free exchange of ideas, treating other people with respect, and promoting diversity, this does not address perhaps a wider and more important issue: the nature of ethical conduct itself in research. Professional ethics in this regard cannot be circumscribed by this or any other code. A code is a starting point in a dialogue that would change the nature of how economists view their own discipline: "Properly understood, professional ethics... is a conversation rather than a constraint, a dance rather than a pose" (DeMartino and McCloskey 2016: 5).

Two significant market failures arise in scientific research in economics: (1) asymmetric information is a problem because investigators know more about their data and procedures than do journal editors or final readers; and (2) replicated research constitutes a public good for society, yet most people free-ride on that discovery process (Wight 2016: 137). Replication of research is rare, even though raw data files and program codes are now required by prominent journals. When replication is attempted, errors are often found. Because there is a noted publication bias for finding statistically significant results, and because lifetime employment through tenure is often on the line when trying to publish, the incentives to cheat can be high, and the costs of cheating low (because detecting and proving cheating is difficult). Moral hazard arises, and the standard economic approach to using incentives may be of limited value in these cases. Reputation effects can reduce cheating only when there is public acknowledgment and shaming of those who transgress; no transparent mechanism exists for this in economics. In a survey of editors at prestigious journals, only half said that they would report cases of plagiarism to an author's supervisors (Enders and Hoover 2004).

In experimental economics, a longstanding concern is over "developing and maintaining a reputation among the student population for honesty" in order to ensure that subjects are responding to an experiment's incentives rather than to anticipated manipulations (Wilson 2016: 317–18). In this regard, experimentalists who deceive research subjects can poison the well of trust, causing negative psychological externalities among potential student subjects. To some degree this problem might be handled by an on-campus institutional review board or by funding agencies that could withhold funding

for the research. Ultimately, the prevention of deception comes down to the issue of human character: “At this point, like in all research, we must rely on the integrity of the experimenter” (ibid.: 320).

Frank Knight described the social milieu within which economics research advances as being akin to a “religion.” Converts adopt an implicit moral code like a monk adopts the habit of self-denial:

Now scientific enquiry has, and rests upon, a moral code, or in sheer fact a “religion”; and it is supremely important that scientists recognize this fact. . . . The basic tenet of scientific research—truth or objectivity—is essentially a moral principle, in opposition to any form of self-interest. (1947: 244)

Honesty is a responsibility to others, not simply a means to enhance one’s own utility. “Honesty is the best policy,” quipped Richard Whately, the nineteenth-century economist-philosopher, “but he who is governed by that maxim is not an honest man” (Ratcliffe 2011: 210). If the motive for honesty is the desire to manipulate outcomes, that is merely opportunistic.

Horizontal pluralism demonstrates how more than one ethical framework helps create trust in the economics profession. Trust lowers transaction costs and enhances the opportunity and value of the exchange of information. A virtuous economist who is motivated by intrinsic convictions about honesty helps produce positive outcomes for the community of scholars and potentially the wider society. But the good outcomes are not what drives the behavior: good economists instinctively abhor dishonesty and are disgusted by its purveyance. Repugnance is the instinctual moral sentiment that provides the basis for norms of behavior within the Code of Conduct proposed by the AEA. Mentoring young colleagues and students through personal example is often the best way of inculcating ethical practices. Samuel Bowles quotes Confucius on this point: “Guide them with government orders, regulate them with penalties, and the people will seek to evade the law and be without shame. Guide them with virtue, regulate them with ritual, and they will have a sense of shame and become upright” (2016: 11). A pluralist account of how and why economists are honest in practice recognizes and bolsters the institutions that promote trustworthiness.

4.5 ARGUMENTS AGAINST PLURALITY

4.5.1 Ultimate Ends

To neoclassical welfare economists, the preference satisfaction approach to understanding human well-being, while not perfect, captures much of the instrumental nature of human desire and action. All human activity can be said to rationally derive from this overarching goal of being preferred, regardless of the packaging used to

pretty it up. I might give \$1,000 to a charity, but my goal is to satisfy my own ends. I might fulfill a duty to another, but purely to satisfy my own ends. A utility function, it is argued, can account not only for selfish desires, but also for altruism and other-regarding preferences that can encompass justice, duties, and even virtues. Satisfying preferences thus is the “one” good thing within which all other values and principles can be subsumed. While Smith and Sen would reject such a claim, a pure neoclassical might argue that ethical pluralism can lead to an operational dead end of relativism that provokes conflict rather than solutions.

Economists might argue, moreover, that they have no expertise in ethics, and the effort to remedy this by introducing pluralism may not be worth the trouble, since the Kaldor-Hicks approach to welfare is acceptable. David Friedman writes, “As an economist, I have no expertise in good and bad. I can, however, set up a ‘criterion of goodness’ called efficiency . . . [Such] an imperfect criterion of desirability is better than none” (1986: 347). One could also claim that economists should not learn new modeling or estimation techniques for the same reason: that existing methods are adequate. But this argument is unconvincing, because one does not know the loss function unless one explores findings in other fields, and economists are always learning and adapting as needed.

Herman Daly, an outspoken neoclassical critic, nevertheless argues that “Ethics is the ordering of multiple ends into a hierarchy with reference to some vision of the Ultimate End, however dimly we perceive it” (2016: 169). To Daly, an “ethical incoherence comes from pluralizing the Ultimate End” (ibid.: 170). Utilitarians also posit an ultimate end (such as pleasure or happiness) that can capture all relevant information, but Daly’s approach is not top-down like a utilitarian’s might be. Rather, because the ultimate end is only vaguely known to ourselves ahead of time, “it is only in the bottom up process of struggling to rank competing ends in specific situations that we get an insight into what the Ultimate End must be like for our consciences to approve the decisions” (ibid.: 169). Without reference to an anchor of an ultimate end, discourse breaks down in fighting, “either with force or deceit” (ibid.: 173).

This account seems to demand a universal acceptance of that ultimate end, but there is no guarantee that any human process would arrive at that. It provides omniscient power to our “consciences” to approve or disapprove of this end, without discussing what a conscience is, where it comes from, or how it might evolve. In this regard, it seems to draw upon intuition or moral feeling without recognizing that this is not an end, but more akin to Adam Smith’s moral sentiments.

4.5.2 Unifying across Frameworks

A different approach, by Derek Parfit, is to attempt to show that various moral frameworks—in his case, utilitarianism, Kantianism, and contractualism—are engaged in the same endeavor, or “climbing the same mountain on different sides” (2011: xxiii). To Parfit, these approaches can be combined into a unifying framework called the Triple Theory: “An act is wrong just when such acts are disallowed by some principle that is optimistic,

uniquely universally willable, and not reasonably rejectable” (ibid.: 413). There is no need for pluralism when all different accounts are really all the same thing under different guises.

While many could agree with Parfit that rule utilitarianism and Kantianism can recommend similar ethical principles, it is not plausible that this would satisfy all objections. For one thing, economists—despite what they may say—are not utilitarian in either spirit or method, and would not fit within the Triple Theory for reasons elaborated earlier. Economists differ from utilitarians along two important dimensions. First, modern welfare economists make no claims about the nature of utility per se; there is no mental state called happiness or pleasure or pain that is of any interest to economists. Second, modern welfare economics does not treat the interests of all sentient beings equally, as in the utilitarian calculus. People are accorded a moral interest in proportion to their spending power, and the interests of the poor are reduced or ignored.

Parfit’s solution to integrating various moral theories does not, ultimately, solve the problem of pluralism because there are more ethical approaches (including the modern welfare economic view) than the Triple Theory can handle. The virtue in Kantian ethics is also substantially different from that of Adam Smith’s ethics, whose moral sentiments intuition for deriving rules or laws is antithetical to Kantian rationalism. These problems are not a flaw but a feature of the complexity of social life. Arjo Klamer, quoted earlier, is an academic economist who also serves as the governor of a Dutch town. The complex practice of reconciling theory with application to social life has led him to a pluralist approach, which he calls “value-based.” Whereas a standard economic approach uses information like preferences, constraints, and prices as inputs into a logical decision making process, in the value-based approach people interact to place values on things through a non-rational progression:

I use the Greek word *phronesis* to characterize the process of valorization. *Phronesis* involves the weighing of values, conditions, interests and findings in order to do the right thing. It is often a chaotic process with a lot of talk, a going back and forth, the making of mistakes, and more talk and deliberation. (2016: 372)

To many this chaotic approach reflects all that is bad about public policy, and could be improved with the economist’s clarifying approach using clear values, methods, and outcomes that would produce a single, obviously correct answer. Such a world may not exist.

4.5.3 Relativism

The complaint that ethical pluralism leads to a dead end of relativism is a concern. How does one navigate in a place of uncertainty within and between moral frameworks? One approach is through pragmatism, treating each case as an opportunity to reflect. McCloskey and Ziliak (2010), for example, discuss how a researcher using econometrics must make informed judgments rather than follow absolute rules, if a truth worth knowing is to be discerned. Imagine a study on the impact of a pollutant

on fetal brain development. Can we prioritize the concern for not accepting a statistical false positive over our concern for the loss function in this instance—that is, the cost of being wrong, and living with the consequences of spina bifida? In such a world it is mistaken to follow simple rules:

Because of the trade-off that may need to occur between statistical significance and practical importance, a bright-line rule of statistical significance would be poor practice. That is, it would be impossible to construct an across-the-board rule that could take into account the case-by-case balancing between practical importance and statistical significance that may be desired. (McCloskey and Ziliak 2010: 13)

Practical importance has to be balanced against other considerations, and this can only be done in a messy way through conversation and debate.

The pluralist account of how to carry out normative economics and public policy does not lend itself to finding the “right” answer that can be arrived at by scientific method alone. Economic efficiency is one goal to be considered in a reflection about the right value and the right method of analysis; a policy solution is reached through considering multiple angles and through rhetorical discourse, which is closer to describing the actual process of truth seeking, according to Ronald Coase (1994: 18). Put differently by Isaiah Berlin, “no perfect solution is, not merely in practice, but in principle, possible in human affairs” (2013: 50). If this is so, then the economist’s concern for maximizing efficiency might give way to a pluralist approach in which efficiency is one of several outcomes, and consequentialism one of several ethical frameworks that should be considered. In thinking about the advantages of such complexity, Kenneth Arrow notes that “looking at policy issues from the point of any one system is likely to lead to unsatisfactory conclusions somewhere. The multiplicity of control systems in the real world is probably no accident” (1997: 765). Each normative framework offers a check on the others, which it would be a mistake to eliminate, even if doing so were philosophically more elegant and offered the hope of absolute ethical answers.

4.6 CONCLUSION

Philosophical coherence may be sorely lacking in muddling through with a pluralist process, yet it is not always clear that coherence should be upheld as the highest standard (Marino 2015: 5). Adam Smith was a pragmatic muddler who, while passionately extolling the principle of liberty, also made numerous exceptions, such as in advocating for interest rate controls to improve capital allocations in financial markets. The restricted liberty of a few bankers (in this case) did not outweigh the benefits to society of paternalistic regulations that would reduce excessive risk-taking and promote stable growth.

There is no simple rule or principle that can satisfy Smith, or that would fit every case, which leaves us always with a suboptimal result with regard to any particular principle

or procedure. Yet muddling through may be the best we can do at a particular point in the social and political conversation. Most policies, therefore, as with most personal ethical decisions, involve an element of deficiency. To Smith this is a feature of progress itself, and “though not the best in itself, it is the best which the interests, prejudices, and temper of the times would admit of.” The future will be rosier if our imperfect actions today “prepare the way for a better” (1776: 543).

ACKNOWLEDGMENTS

The author gratefully acknowledges valuable suggestions by Scott Davis and Mark White.

REFERENCES

- AEA. 2018. “Draft Code of Professional Conduct.” Available at <https://www.aeaweb.org/resources/member-docs/draft-code-of-conduct>.
- Arrow, Kenneth J. 1963. “Uncertainty and the Welfare Economics of Medical Care.” *American Economic Review* 53: 941–73.
- Arrow, Kenneth J. 1997. “Invaluable Goods.” *Journal of Economic Literature* 35: 757–65.
- BBC News. 2017. “Passenger Dragged Off United Airlines Plane Wins Settlement.” April 27. Available at <http://www.bbc.com/news/world-us-canada-39739737>.
- Becker, Gary S, and Julio Jorge Elias. 2007. “Introducing Incentives in the Market for Live and Cadaveric Organ Donations.” *Journal of Economic Perspectives* 21(3): 3–24.
- Berlin, Isaiah. 2013. *The Crooked Timber of Humanity: Chapters in the History of Ideas*, 2nd edn. Henry Hardy (ed.). Princeton, NJ: Princeton University Press.
- Bowles, Samuel. 2016. *The Moral Economy: Why Good Incentives Are No Substitute for Good Citizens*. New Haven, CT: Yale University Press.
- Bowles, Samuel, and Herbert Gintis. 2011. *A Cooperative Species: Human Reciprocity and Its Evolution*. Oxford: Oxford University Press.
- Brecht, Bertolt. 1941. *Mother Courage and Her Children*. Eric Bentley (trans.). New York: Grove Press (1955 edition).
- Campbell, John. 2018. “Interim Report of the Ad Hoc Committee to Consider a Code of Professional Conduct.” Available at <https://assets.aeaweb.org/assets/production/files/6219.pdf>.
- Coase, Ronald H. 1994. *Essays on Economics and Economists*. Chicago, IL: University of Chicago Press.
- Daly, Herman. 2016. “Ethics in Relation to Economics, Ecology, and Eschatology.” In George DeMartino and Deirdre McCloskey (eds.), *The Oxford Handbook of Professional Economic Ethics* (Oxford: Oxford University Press), pp. 169–83.
- DeMartino, George F. 2011. *The Economist’s Oath: On the Need for and Content of Professional Economic Ethics*. New York: Oxford University Press.
- DeMartino, George F. and Deirdre N. McCloskey (eds.). 2016. *The Oxford Handbook of Professional Economic Ethics*. New York: Oxford University Press.
- DeMartino, George F. and Deirdre N. McCloskey. 2018. “Professional Ethics 101: A Reply to Anne Krueger’s Review of *The Oxford Handbook of Professional Economic Ethics*.” *Economic Journal Watch* 15: 4–19.
- Enders, Walter, and Gary A. Hoover. 2004. “Whose Line Is It? Plagiarism in Economics.” *Journal of Economic Literature* 42: 487–93.

- Folbre, Nancy. 2012. *For Love and Money: Care Provision in the United States*. New York: Russell Sage Foundation.
- Frank, Robert H., Thomas Gilovich, and Dennis T. Regan. 1993. "Does Studying Economics Inhibit Cooperation?" *Journal of Economic Perspectives* 7: 159–71.
- Frey, Bruno S. and Felix Oberholzer-Gee. 1997. "The Cost of Price Incentives: An Empirical Analysis of Motivation Crowding-Out." *American Economic Review* 87: 746–55.
- Friedman, David D. 1986. *Price Theory*. Cincinnati, OH: South-Western.
- Friedman, Milton. 1962. *Capitalism and Freedom*. Chicago, IL: University of Chicago Press.
- Friedman, Milton. 1970. "The Social Responsibility of Business Is To Increase Its Profits." *New York Times Sunday Magazine*, September 13.
- Gintis, Herbert, Samuel Bowles, Robert Boyd, and Ernst Fehr (eds.). 2005. *Moral Sentiments and Material Interests: The Foundations of Cooperation in Economic Life*. Cambridge, MA: MIT Press.
- Gneezy, Uri, Stephan Meier, and Pedro Rey-Biel. 2011. "When and Why Incentives (Don't) Work to Modify Behavior." *Journal of Economic Perspectives* 25: 191–210.
- Gould, Skye, and Dave Mosher. 2017. "Americans Spent \$8 Billion on Plastic Surgery in 2016—Here's the Work They Got Done." *Business Insider*, May 22. Available at <http://www.businessinsider.com/plastic-surgery-growth-statistics-facts-2016-2017-5>.
- Klamer, Arjo. 2016. "The Value-Based Approach to Cultural Economics." *Journal of Cultural Economics* 40: 365–73.
- Knight, Frank H. 1947. *Freedom and Reform*. New York: Harper and Brothers.
- Knight, Frank H. 1960. "Social Economic Policy." *Canadian Journal of Economics and Political Science* 26: 19–34.
- Kochanek, Kenneth D., Sherry L. Murphy, Jiaquan Xu, and Elizabeth Arias. 2017. "Mortality in the United States, 2016." NCHS Data Brief No. 293. Available at <https://www.cdc.gov/nchs/products/databriefs/db293.htm>.
- Levy, David M., and Sandra J. Peart. 2016. *Escape from Democracy: The Role of Experts and the Public in Economic Policy*. New York: Cambridge University Press.
- Marino, Patricia. 2015. *Moral Reasoning in a Pluralistic World*. Montreal: McGill-Queen's University Press.
- McCabe, Kevin A., Mary L. Rigdon, and Vernon L. Smith. 2003. "Positive Reciprocity and Intentions in Trust Games." *Journal of Economic Behavior and Organization* 52: 267–75.
- McCloskey, Deirdre N., and Stephen T. Ziliak. 2010. "Brief of Amici Curiae Statistics Experts in Support of Respondents." Vol. No. 09–1156. Washington, DC: Supreme Court of the United States. Available at https://www.americanbar.org/content/dam/aba/publishing/preview/publiced_preview_briefs_pdfs_09_10_09_1156_RespondentAmCu2Profs.authcheckdam.pdf.
- Mill, John Stuart. 1848. *Principles of Political Economy with Some of Their Applications to Social Philosophy*. Stephen Nathanson (ed.). Indianapolis, IN: Hackett (2004 edition).
- Mill, John Stuart. 1863. *Utilitarianism*. George Sher (ed.). Indianapolis, IN: Hackett (2002 edition).
- Nelson, Julie A. 2016. "Poisoning the Well, or How Economic Theory Damages." In George DeMartino and Deirdre McCloskey (eds.), *The Oxford Handbook of Professional Economic Ethics* (Oxford: Oxford University Press), pp. 184–99.
- Ratcliffe, Susan (ed.). 2011. *Oxford Treasury of Sayings and Quotations*, 4th edition. Oxford: Oxford University Press.
- Pareto, Vilfredo. 1906. *Manual of Political Economy*. Ann S. Schwieger (trans.), Ann S. Schwieger and Alfred N. Page (eds.). New York: Augustus M. Kelley (1971 edition).

- Parfit, Derek. 2011. *On What Matters: Volume I*. Oxford: Oxford University Press.
- Reinhardt, Uwe E. 2001. "Can Efficiency in Health Care Be Left to the Market?" *Journal of Health Politics, Policy and Law* 26: 967–92.
- Rosenberg, John S. 2001. "Toxic Memo." *Harvard Magazine*, May–June. Available at <https://harvardmagazine.com/2001/05/toxic-memo.html>.
- Sen, Amartya K. 1970. *Collective Choice and Social Welfare*. San Francisco, CA: Holden-Day.
- Sen, Amartya K. 1977. "Rational Fools: A Critique of the Behavioral Foundations of Economic Theory." *Philosophy & Public Affairs* 6: 317–44.
- Sen, Amartya K. 1987. *On Ethics and Economics*. Oxford: Blackwell.
- Sen, Amartya K. 1999. *Development as Freedom*. New York: Knopf.
- Smith, Adam. 1759. *The Theory of Moral Sentiments*. D. D. Raphael and A. L. Macfie (eds.). Indianapolis, IN: Liberty Fund Press (1982 edition).
- Smith, Adam. 1776. *An Inquiry into the Nature and Causes of the Wealth of Nations*, R.H. Campbell and A. S. Skinner (eds.). Indianapolis, IN: Liberty Press (1981 edition).
- Smith, Vernon L. 2013. "Adam Smith: From Propriety and Sentiments to Property and Wealth." *Forum for Social Economics* 42: 283–97.
- Stiglitz, Joseph E, Amartya Sen, and Jean-Paul Fitoussi. 2009. *Report by the Commission on the Measurement of Economic Performance and Social Progress*. Available at <http://ec.europa.eu/eurostat/documents/118025/118123/Fitoussi+Commission+report>.
- Wight, Jonathan B. 2014. "Economics within a Pluralist Ethical Tradition." *Review of Social Economy* 72: 417–35.
- Wight, Jonathan B. 2015. *Ethics in Economics: An Introduction to Moral Frameworks*. Stanford, CA: Stanford University Press.
- Wight, Jonathan B. 2016. "The Ethical Economist: Duty and Virtue in the Scientific Process." In George DeMartino and Deirdre McCloskey (eds.), *The Oxford Handbook of Professional Economic Ethics* (Oxford: Oxford University Press), pp. 137–64.
- Wight, Jonathan B. 2017. "The Ethics behind Efficiency." *Journal of Economic Education* 28: 15–26.
- Wilson, Bart J. 2016. "The Meaning of *Deceive* in Experimental Economics." In George DeMartino and Deirdre McCloskey (eds.), *The Oxford Handbook of Professional Economic Ethics* (Oxford: Oxford University Press), pp. 317–28.
- Wu, Alice H. 2017. "Gender Stereotyping in Academia: Evidence from Economics Job Market Rumors Forum." Manuscript.
- Zak, Paul J. 2009. "The Physiology of Moral Sentiments." *Journal of Economic Behavior and Organization* 77: 53–65.

CHAPTER 5

ECONOMIC ETHICS AND THE CAPABILITY APPROACH

CONSTANZE BINDER AND
INGRID ROBEYNS

5.1 INTRODUCTION

IN the last decade, scholarly interest in the capability approach as a normative framework has exploded. Scholars from a broad variety of disciplines are using this framework to conduct various types of normative analyses, whether evaluative, prescriptive, critical or as the basis for policy proposals and institutional design. The capability approach is generally understood as a normative framework, although it can in principle also be used as providing tools for descriptive or explanatory research (Robeyns 2017: 142–3). As a consequence of its normative orientation, it should not be surprising that the capability approach is also a relevant framework for economic ethics.

Moreover, the main contemporary scholar who has introduced the capability approach in its current, explicit formulation was Amartya Sen in his 1979 Tanner Lecture (published as Sen 1980). In that lecture Sen asked the famous “equality of what?” question: assuming we advocate a form of equality, what kind of good should be equalized? In other words, what should be the metric or “currency” (Cohen 1989) of egalitarianism? The introduction of the idea of *capability* in the literature on equality and theories of social and distributive justice led to a large literature debating various aspects of capabilitarian theories of justice. Several philosophers proposed more fully worked-out capabilitarian theories of justice. Martha Nussbaum’s capability theory of social justice (2000, 2006) has been the most influential, in part because it contains a much-cited list of which capabilities a political community should guarantee to its citizens. Yet a wide range of capabilitarian theories of justice are possible, including the more perfectionist theory

defended by Richard Arneson (2010) or the agency-based theory developed by Rutger Claassen (2017).

While Sen coined the idea of capability in the specific discussion on equality, it soon became understood as a proposal within several other normative theories. In the following years, Sen introduced the capability approach also in debates about the measurement of well-being in developing countries (Sen and Sengupta 1983; Sen 1985a), and questions about conceptions of well-being and freedom (Sen 1984, 1985b). Given the centrality of those notions to economic ethics, the capability approach is a tool that could potentially be used for analysis in economic ethics. How such capability analyses in economic ethics would look like, and what their promises and limitations are, is the topic of this chapter.

Since the term “economic ethics” is not widely used among ethicists working on economic questions or among economists interested in ethical questions, we should first introduce it. We use the term as a shorthand for “normative analyses in social and political philosophy and ethics that focus on economic institutions, systems, and practices.” Although the term “economic ethics” may not be so widely used as similar terms in other fields of applied ethics (such as “medical ethics,” “bioethics,” or “ethics of technology”), we believe that the recent resurgence of normative philosophy on economic topics justifies a more widespread use of this term. Having a clear term could, possibly, also help this field to grow and become stronger—a goal for which the capability approach also has a potential role to fulfill, as we will show.

This chapter will give an account of the capability approach and then discuss its contribution to economic ethics. In Section 5.2, we describe the capability approach and focus in more depth on those aspects that are relevant for economics and ethics. The next two sections examine in some detail the two core notions in the capability approach: freedom (Section 5.3) and well-being (Section 5.4). This account allows us to ask how the capability approach can be (and has been) used in economic ethics, as well as to point at some unresolved issues that capability scholars need to pay attention to if they want to put the capability approach to use when analyzing questions of economic ethics.

5.2 WHAT IS THE CAPABILITY APPROACH?

If one had to summarize the capability approach in one sentence, one might say that it is a theoretical framework urging us, when asking normative questions, to ask what people are able to do and what lives they are able to lead. The capability approach is concerned with aspects of people’s lives such as their health, the education they can enjoy, and the support they enjoy from their social networks; it is also concerned with what people can do, such as being able to work, raise a family, travel, or be politically active. The capability approach cares about people’s real freedoms to do these things, and the level of well-being that they will reach when choosing from the options open to them. It is a rich, multidimensional approach.

The capability approach is a framework which entails specific conceptualizations of human well-being and freedom, which in turn can be used in the study of related concepts, such as poverty or human development. A person's well-being and related concepts are assessed in terms of the "doings and beings" a person can realize in her life. A person can derive different doings and beings, her so-called *functionings*, from the resources at her disposal. All those doings and beings that she can realize, given her resource endowment and the physical, social, and environmental aspects that affect the conversion of resources into doings and beings, are included in her so-called *capability set*. Those functionings she can jointly realize are called *bundles of functionings* and those that she actually chooses to realize are referred to as *achieved functionings*.¹

Dependent on the problem it is meant to address and the conventions of the discipline in which it is used, the general capability framework can be filled out in different ways, leading to different specifications. When studying or applying the capability approach, it is very important to make the distinction between the *capability approach*, which is the general framework that is used in a number of different disciplines for various purposes, and *capability theories* or *capability accounts*, which specify the general framework in order to use it for a particular purpose, such as to develop a theory of well-being, poverty, or justice (Robeyns 2016a, 2017: 29–30). This distinction is important, because a claim or an argument made by a capability scholar could be either a claim about the capability approach in general, or about particular capabilitarian theories, that are developed in a particular debate (such as theories of distributive justice, development ethics, or welfare economics) (Binder 2009). In order to properly understand a particular capabilitarian claim or argument, one needs to know what the exact scope of the claim is. Elsewhere we have argued at length what all capabilitarian theories share (Robeyns 2017). Let us here highlight three core characteristics shared by all capabilitarian theories that are particularly important in order to understand their contribution to economic ethics.

A first important commitment of all capabilitarian theories is the shift of the evaluative space for the assessment of well-being and related concepts, such as development or collective welfare, to a person's capabilities and functionings. Functionings are all doings and beings; one way to identify those relevant for a particular evaluative exercise is to identify all those the person in question has reason to value. A person's capability set then includes all those functioning bundles at her command, given her resource endowment and her personal, social, or environmental characteristics. These characteristics entail her set of "conversion factors," those that determine the degree in which a person can transform a

¹ Note that the definition of functionings and capability slightly differs in the literature. We opt in this chapter to follow Sen's earlier notation due to its close connection to the literature on overall freedom in political philosophy and freedom-rankings in welfare economics. It is important to note, though, that Nussbaum as well as wide parts of the capability literature use the term *capabilities* to refer to all doings and beings a person can potentially realize. In the notation employed in this chapter, those would correspond to all bundles of functionings in a person's capability set. Similarly, Nussbaum uses the term *functionings* to refer to those doings and beings a person does in fact realise, which in this chapter are referred to as realized functionings. For a more detailed discussion of the differences in notation, see Robeyns (2017: 90–8).

resource into a functioning. What is crucially important is that persons have different abilities to convert resources into functionings, and hence a welfare analysis focusing merely on means or resources will not capture the different levels of welfare or well-being that those resources enable to very diverse people (Sen 1992: 19–21, 26–30, 37–8).

A second important commitment of all capabilitarian theories is the move from means to the ends of well-being that comes with the focus on functionings and capability. Instead of focusing on the resources at a person's command, valued for being a means to achieve what the person has reason to value in life, the focus of evaluation is directly and primarily on the ends of a person's well-being: the freedom to achieve the doings and beings a person has reason to value, as reflected by her capability set. The amount and type of resources a person requires to realize certain doings and beings can differ between different people and the societies they live in.

There are two important reasons why the capability approach urges us that we have to start our analysis from the ends rather than the means (Robeyns 2017: 48–9). A first reason is given by the notion of conversion factors: people differ in their ability to convert means into valuable opportunities (reflected by her capability set) or outcomes (her achieved functionings). Because ends are what ultimately matter when thinking about well-being and the quality of life, means can only work as fully reliable proxies of people's opportunities to achieve those ends if all people have the same capacities or powers to convert those means into equal capability sets. However, the capability approach (rightly) assumes that there are very significant differences between individuals in the conversion of resources into functionings. The second reason why the capability approach requires us to start from ends rather than means is that there are some vitally important ends that do not depend very much on material means, and hence would not be picked up in our analysis if we were to focus on means only. For example, self-respect, supportive relationships in school or in the workplace, and friendship are all very important ends that people may value and desire; yet there are no crucial means to those ends that one could use as a readily measurable proxy. We need to focus on ends directly if we want to capture what is important.

A third important commitment of all capabilitarian theories is its commitment to pluralism, including not only pluralism of dimensions in its accounts of well-being, but also value pluralism at the level of what matters in evaluations. The first type of pluralism is also captured by the term "multidimensional analysis," and essentially posits that the key notions in the capability approach—well-being and freedom—and those concepts that build on those notions—such as poverty, riches, inequality, and distributive justice—can never be reduced to a single dimension, but will consist of a plurality of dimensions, such as being healthy, being educated, enjoying basic liberties, and so forth. The second type of pluralism entails that the capability approach does not claim that well-being and freedom are the only things that matter; instead, often there are other important values and moral principles that are equally (or more) important, and that may sometimes more fruitfully be conceptualized *outside* the capabilitarian framework. For example, process values such as transparency or the absence of corruption are not best captured in a capability language, but are nevertheless important.

When compared with other normative approaches, the capability approach distinguishes itself from the attention it gives to *both* freedom and well-being, rather than giving exclusive attention to only one of them and denying the ultimate importance of the other, as well as by the specific conceptualizations of freedom and well-being. We will therefore now turn to the role of freedom (Section 5.3) and the role of well-being in the capability approach (Section 5.4).

5.3 FREEDOM AND THE CAPABILITY APPROACH

One characterizing feature of the capability approach is the importance it ascribes to human freedom. In this section we discuss the main characteristics of the conception of freedom used in the capability approach and the role assigned to it in the assessment of well-being and related concepts. In Section 5.5 we will then see that the specific conception of freedom employed in the capability approach allows it to serve as a framework to assess different economic institutions and systems in economic ethics and potentially make a crucial difference in their evaluation compared to other notions of freedom.

In the light of the large number of conceptions of freedom in the philosophical literature, it is useful to have a systematic way to compare them and see what the crucial (conceptual) differences are between various conceptions. One widely accepted general concept of freedom that serves this purpose very well is MacCallum's (1967) triadic formula. MacCallum argues that every conception of freedom goes back to one and the same concept, namely (a) a person (or group of persons), who (b) is free from a set of relevant constraints (c) to do, not do, become, or not become something. Large parts of the philosophical literature on the conception of freedom can be seen as a debate on how these three aspects should be specified. In the following we shall shortly discuss the characterizing features of these aspects in the capability approach.

(a) The main focus in the capability approach is on individual human beings. The question whether this can and should be extended to group functioning and capability, reflecting the freedom of groups or collectives, is part of a discussion in the capability literature (Ibrahim 2017; Binder and Binder 2017; Robeyns 2017: 115–18). However, the primary focus in the capability framework is on the well-being of individuals. A person's membership in groups can enter the framework due to its importance to certain individual functionings, such as "having access to friends" or as a way of how a person's functionings are facilitated or constrained. An example of the former is the way how group action can enhance certain functionings (such as "speaking a language"). An example of the latter can be group structures that constrain a person's (conversion into) functionings, such as in the case of certain group norms.

(b) The set of relevant constraints is one of the crucial differences to other notions of freedom discussed in the philosophical literature. In the capability framework

outlined by Sen (1985a), it becomes apparent that not only interference by the state or other people are considered to be relevant constraints (as in many conceptions of negative freedom), but also constraints that result from a lack of resources or social and environmental factors that hinder people to realize certain functionings. These correspond to the conversion factors discussed in Section 5.2. What matters is the effective freedom a person enjoys to pursue life paths she has reason to value. These life paths can be blocked by a lack of material resources, factors internal to the person (physical conversion factors), legal rules or social norms (social conversion factors), or even environmental constraints (environmental conversion factors). As opposed to many negative conceptions of freedom that would consider, for instance, a person who cannot afford the money for the school books to be free to go to school (as long as no other person actively interferes with her doing so), according to the notion of freedom employed in the capability approach a person would be unfree to pursue the functioning “going to school” if she lacks the money to do so. Similarly, a social norm that prohibits girls from going to school would diminish the freedom to pursue the functioning of going to school in the capability framework but not according to conceptions of freedom in which social pressure (without actual interference) does not count as a constraint to one’s freedom.

The crucial difference compared to other conceptions of freedom is that societal aspects or material resources can constrain a person’s freedom as well, as reflected by her capability set. What matters is the real or effective freedom a person has to pursue a certain life path, reflected by a bundle of functionings open to her (taking all things into account). This notion of freedom that accounts for a wider set of possible constraints that can hinder people to pursue and realize life-paths they value was termed “effective freedom” by Sen (1992: 36–8, 64–6). In Section 5.5 we will see that this feature, amongst others, allows the capability approach to play a crucial role in the evaluation of economic institutions and systems.

(c) The question as to which doings or beings are considered to be relevant freedoms and thus contribute to a person’s overall freedom is a highly debated issue in the literature on overall freedom. Carter (1999) prominently argued for the need for a value-neutral approach to overall freedom in order to account for freedom’s non-specific value. Sen (2002, 2009), on the other side, argued for the need to account for a different value of freedom, namely for the value and role that effective freedom, as reflected by a person’s capability set, has for a person’s well-being. In order to capture this role of freedom, the value assigned to the different doings and beings open to a person needs to be taken into account. For this purpose, it is required to weight different doings and beings in light of the reasons people have to value them. The question about the valuation of functionings is closely related to the debate within the capability literature as to how the relevant functionings that contribute to a person’s capability or effective freedom should be chosen. Even though there are different stances in the capability literature regarding how the value of functionings should be identified, there is broad agreement that their value, identified by either a list of valuable dimensions (Nussbaum 2000, 2006, 2011) or a process of discussion and deliberation among the people concerned (Sen 2004;

Crocker 2008; Byskov 2017), should play a crucial role in assessing their importance to a person's overall capability expansion.

The reasons that the value of the doings and beings open to a person play a crucial role in the conception of freedom employed in the capability approach become even clearer if one has a closer look at the reasons that freedom is valued in the capability approach. It is valued for at least three different reasons (Sen 1992). First, it adds information when evaluating a person's achieved well-being. To illustrate by drawing on a prominent example from Sen, if we assess the well-being of a person who is not well-nourished, it will be crucial to know whether that person had the freedom to eat but chose to fast for instance for religious or political reasons, or whether the person did not have the option to eat open to her and is thus starving. Second, freedom is valued from the perspective of a policymaker in a plural society. Freedom is important in the assessments of social states because it allows for a plurality of conceptions of the good life in one's society. Moreover, having a variety of valuable life paths open allows people to reflect on their conception of the good in the light of the conception of others in their society, and possibly revise and change their goals and objectives in life in the light of it. Third, another reason to assign value to freedom in the assessment of a person's well-being and social states is the value freedom itself can have to a person's well-being. It might be that a person assigns intrinsic value to freedom, allowing her to deliberate and choose among those paths she has reason to value before embarking on them. It is important to note, however, that the use of the general capability framework is not restricted to notions of well-being that assign such intrinsic value to freedom; its generality can accommodate it but does not presuppose it.

In the light of all these reasons that freedom is considered valuable in the capability approach, it becomes apparent that the nature and value of an option (that is, of a bundle of functionings) plays a crucial role in the question as to whether the respective bundle adds to the expansion of a person's capability set and thus to the expansion of her effective freedom. According to the first reason, it is not any alternative option that will allow a person to fast instead of to starve, but the specific option of "eating" that needs to be available. In case of the second reason, it is not any bundle of functionings a policymaker should account for in a plural society, but rather the bundles of functionings people with different conceptions of the good living in that society value. Last but certainly not least, if freedom has intrinsic value for a person's well-being, then the value of bundles of functionings will be taken into account in accordance with the reasons a person has to value them. One of these reasons is precisely the intrinsic value one ascribes to freedom of choice (and thus to any further choice option).

To conclude this section, we sum up some of the characterizing and distinguishing features of the conception of freedom as employed in the capability approach. These characterizing features will then serve in Section 5.5 to explore and discuss the suitability of the capability approach to assess economic institutions and systems. On the level of relevant constraints a number of distinguishing characteristics can be identified. First, the set of relevant constraints is larger than those of many other notions of freedom. Beside legal constraints, the capability approach also accounts for economic

and non-economic constraining influences, such as a person's resource endowment or her social, psychological, cultural, and environmental conditions. Second, the different constraints can be overcome in different ways and can sometimes partly be compensated for by additional resources or lifting other constraints. Suppose a person is paralyzed and the functioning "moving around" is not open to her due to a physical constraint. In this case additional resources, such as a wheelchair, and the lifting of certain physical constraints in the person's surrounding, such as wheelchair access to public buildings, can (partly) lift the constraint and make the functioning "moving around" available to this person. Third, dependent on the respective constraint, different actors or institutions might be in the best position to lift the respective constraint and contribute to the expansion of a person's capability. In many cases this can be the government by lifting or imposing certain legal constraints or guaranteeing some functioning related to health or education.² In other cases, however, it can be an NGO or a biomedical or technical innovation that expands a person's capability set, such as in the case of the functioning "living in a malaria-free world." This allows the capability framework to assess different institutions and actors in terms of the role they play in expanding (or limiting) human capability.

Last, but certainly not least, on the level of the alternatives, another important characterizing feature is the importance ascribed to the value of the doings and beings a person is constrained (or not) in pursuing. The value assigned to functionings plays a crucial role when judging as to whether additional functionings open to a person do indeed increase a person's freedom, as reflected by her capability set. More specifically, only those doings or beings that are valued as an end—but not those that are merely valuable as a means—do in fact increase a person's capability.³ In Section 5.5 we will see that this feature plays a crucial role in the assessment of economic systems in terms of their resource use (to achieve a certain level of well-being). Given these characterizing features, one crucial question is whether and if so which difference this expansion of the informational space toward effective freedom plays for economic ethics. More specifically, we ask which difference the capability conception of well-being and freedom makes in the evaluation of economic policies and economic systems. We will take these questions up in Sections 5.4 and 5.5 respectively.

5.4 WELL-BEING AND THE CAPABILITY APPROACH

In our exposition of the capability approach, we have explained that there are two core notions in the capability approach, freedom and well-being. But what, exactly, is the nature of the account of well-being in the capability approach? The first thing to note

² For a discussion on how new rules can lead to an expansion of freedom or capabilities, see Section 5.5.

³ For a discussion about the difficulty to distinguish between functionings that are valued as a means (possibly to bring about other functionings) and as ends in themselves, see Robeyns (2016b).

is that the capability approach includes a notion of *achieved* well-being, focused on functionings, as well as a notion of well-being *freedom*, represented by one's capability set (Sen 1985b, 1993a). The distinction between achieved well-being and well-being freedom is virtually absent from the well-being literature. In contemporary philosophy, most accounts focus on how well life is going for a person—hence, on achieved well-being. But clearly, for policy purposes, we often focus on well-being freedom, because other values, such as respect for personal autonomy or even human dignity, may prevent us from having a *specific* well-being outcome as a legitimate policy goal. When in the capability approach the term “well-being freedom” is used, it refers to what philosophers elsewhere would call “opportunities for well-being.” This notion is especially relevant in moral theories in which we try to balance a concern for well-being with a concern for individual freedom to choose: the term “well-being freedom” tries to bring together and integrate those two values.

The second thing to note is that there are at least two defining characteristics of the capability account of well-being. First, it is multidimensional, and incommensurability between dimensions is generally assumed. For example, health, being educated, being sheltered, engaging in meaningful work, and so forth, are all very different dimensions, and we cannot assume that an increase in one of those dimensions will compensate the well-being loss for a decrease in another dimension. Second, functionings entail both material and nonmaterial dimensions, and also opportunities or outcomes related to the market and the nonmarket spheres. This makes the capability approach an attractive account of well-being for economists who have been arguing that mainstream economics tends to focus exclusively on the outcomes generated by markets. A purely market-based understanding of economics is problematic not only because much economic activity is performed outside the market, but also because non-market economic activities may lead to dimensions of well-being that are not generated (as much or at all) by the market, such as the quality of social affiliation or mental and psychological well-being for vulnerable persons (Folbre 2001).

What kind of well-being account does the capability approach offer? In philosophy, the starting point for answering that question is generally the typology offered by Derek Parfit, who suggests that we should make a distinction between three types of philosophical theories of well-being:

On *Hedonistic Theories*, what would be best for someone is what would make his life happiest. On *Desire-Fulfilment Theories*, what would be best for someone is what, throughout this life, would best fulfil his desires. On *Objective List Theories*, certain things are good or bad for us, whether or not we want to have the good things, or to avoid the bad things. (1984: 493)

The capability approach is often categorized as being an objective list theory (Angner 2015), because functionings are plural and the selection of dimensions gives us a list of items which are judged to be valuable for persons (Qizilbash 2013). However, elsewhere we have argued that there is not merely one well-being account in the capability approach,

but several (Robeyns 2017: 125–6). The reason is that there is a variety of capability theories in the general capability literature, and those theories can employ different accounts of well-being dependent on how they are specified and for which purpose they are constructed.

If a capability theory is used for a first-person perspective, for example by an adolescent contemplating what to do with her life, she may ask herself what she really wants: to study hard and work hard and become a medical doctor? Or does she have a stronger desire to build a family and search for a job that makes it possible to spend enough time with her children? Does she want to devote her life to fighting for a good cause? In this personal deliberation, the account of well-being she then uses can be seen as a desire-fulfillment account in which the relevant functionings are identified on the basis of her desires. In short, the account of well-being used for a specific capability theory depends on the way the relevant functionings are selected.

Which account of well-being in the capability approach is relevant for economic ethics? We take it that economic ethics focuses primarily on the ethical analysis of economic institutions and systems. A helpful way to approach this question is to draw on the distinction between micro-level or macro-level analysis. In the design of policies or institutions on a micro scale, such as in the case of project assessments, it will be possible and often desirable to employ a subjective account of well-being, based on a discussion and deliberation of the people concerned about the doings and beings they value in life and why they do so. Note that some might value a certain functioning due to her desires for it, while others might value a functioning due to the happiness it yields or for other reasons. What matters is that the underlying account of well-being is a subjective one based on a deliberation among the people concerned. Alternatively, policies can have the objective to provide the resources (money and sometimes time) that are means to realize a wide range of functionings different people might value. For policymaking we often have to choose either an approach that uses resources as a proxy for well-being (although this cannot account for differences in conversion factors between people) or else policymakers will try to provide a range of options to us, where ideally the policymaker assumes that these options are things that many people want.

If a capability theory is made for macro-level poverty analysis, such as in the OECD's CLIO-Infra project (Van Zanden et al. 2014) in which economic historians used a broader notion of well-being to make long term historical analyses, the researchers will select a number of functionings that they have reason to believe are good for people, such as their health, educational outcomes, and the kind of shelter in which they can live. The notion of achieved well-being entailed in this normative exercise is an objective account. One common way to justify the adoption of such an objective account is to argue that one has reason to assume that these are dimensions of the quality of life that people would want for themselves and that, given that one is working with very large numbers, it is a safe assumption to proceed this way.

We have now laid out the conceptualizations of freedom and well-being in the capability approach. This will allow us, in the following section, to explore the role of the capability approach in the assessment of economic systems and institutions.

5.5 THE CAPABILITY APPROACH IN ECONOMIC ETHICS

An important part of economic ethics is the assessment of economic and social systems in terms of different values or objectives. One obvious question in this respect is whether and if so how the capability approach would be suitable for this purpose. As discussed in Section 5.2, the capability framework is a broad and open framework that needs to be specified in different ways dependent on the problem at hand or the specific normative stance adopted. To give some examples, the capability approach can lend itself to assess institutional or governmental structures in terms of justice. Nussbaum's (2006) theory is a prominent example of such a capability theory of justice. Indeed, many other areas of the literature focus on the assessment of governmental structures or individual policies (including those undertaken by NGOs, for instance). What is often overlooked, though, is that the capability framework lends itself to assess societal and economic systems in more general ways as well.

In Section 5.5.1, we shall discuss how the characterizing features of the capability approach, discussed in Sections 5.2 and 5.3, are particularly well-suited to assess economic and societal systems. For this purpose, we discuss a number of aspects that all variants of the capability approach share by virtue of being part of the core characteristics of the approach.⁴ We then move, in Section 5.5.2, to a discussion of a number of unresolved questions that need to be addressed in the specification of the capability approach before it can be employed in the assessment of economic systems.

5.5.1 How Does the Capability Approach Make a Difference in Assessing Economic Institutions and Systems?

When it comes to the debate about the comparative merits of different economic systems, to our knowledge the capability approach has hardly been employed so far.⁵ However, the notion of freedom itself has been extensively employed in defense of capitalism. Often a libertarian or negative conception of freedom is used for this purpose. What characterizes both conceptions of freedom is that government interference in one's justly acquired property rights (in case of libertarian conceptions such as Nozick 1974) or legal constraints imposed by the government (in case of some notions of negative freedom such as Hayek 1960 and Oppenheim 1985) limit a person's freedom according

⁴ It is important to note, however, that the actual assessment will obviously depend on the specific objective pursued, such as whether one is interested in a general capability expansion (not necessarily equally distributed among different people in society) or in reaching a sufficient level of capabilities for everyone, in equality of capability at the beginning of one's life (possibly leading to unequal capabilities at later points in time due to people's choices) or equality of capability throughout people's lives.

⁵ For a notable exception, see Claassen (2009).

to the respective notion of the concept. A lack of resources on the other hand, such as that due to an unequal distribution of resources, does not necessarily count as a constraint of a person's freedom. As a result, in these conceptions of freedom often used by defenders of capitalism, the absence of state interference in the form of rules, regulations, or taxation leads to an increase in freedom and thus favors the free market (often equated with capitalism). A lack or abundance of resources, on the other hand, cannot constrain or enhance a person's freedom according to these conceptions of freedom.

Alternatively, those who defend socialism in terms of freedom often adopt a so-called positive notion of freedom in which provision by the government of certain facilities counts as expanding a person's freedom (Cohen 1981). According to many positive conceptions of freedom, however, direct or indirect interference by the government does not count as an infringement upon a person's freedom. Indeed, prominent critics pointed to the risk that some positive conceptions of freedom would justify government manipulation in the name of freedom (Berlin 1969). As a result, in the literature that discusses the assessment of markets, defenders of markets often employ a negative notion of freedom, while critics of markets tend to employ a positive conception of freedom. This often leads to a stalemate in which one's position depends on the notion of freedom adopted.

The notion of freedom employed in the capability approach promises to move beyond this dichotomy in the literature. As detailed in Section 5.3, within the notion of freedom employed in the capability approach, a lack of resources can limit a person's capability set, as does a legal rule that prohibits a person to exercise certain (negative) rights. This allows us to move beyond the dichotomy present in the literature. In the case of a capability assessment of societal systems, a lack of resources due to unequal distribution generated by a capitalist market can constrain a person's capability set, and so can the interference by a socialist government that limits (for example) a person's freedom to worship. Similarly, the generated wealth in a capitalist market system can expand one's capability set as much as the state provision of education and health care in a socialist system.

The large set of constraints considered to be relevant for the conception of "effective freedom" as employed in the capability approach allows us to assess both the merits and risks of capitalist and socialist systems by considering all possible ways these can lead to a limitation or expansion of a person's capability set, thereby enabling a more nuanced assessment of their respective merits and drawbacks. It is precisely this point that allows one to move beyond the debate about "socialism versus capitalism" to the development and assessment of a new societal system that combines the market mechanism and governmental structures and institutions in a way that leads to an expansion of human capability. Thus, the first characterizing feature of the conception of freedom employed in the capability approach, namely that it accounts for a wide set of constraints limiting a person's effective freedom, allows one to assess both the merits and pitfalls of capitalist and socialist systems alike and thus provide a stepping stone in moving beyond the dichotomy between them and toward the study and development of new economic systems.

Another important feature of this wide set of relevant constraints that includes a lack of resources as an obstacle to a person's freedom is the general notion of resources employed in the capability approach. As already discussed in Section 5.4, the framework is not restricted to commodities traded on markets, but can also include resources that were produced in a subsistence economy or in the household, as well as non-material resources such as time availability and human capital. This feature allows the capability approach to assess market systems as well as informal economic systems such as subsistence farming or intra-household production.

A third characterizing feature that we discussed in Section 5.3 is the possibility of considering conversion factors as either imposing or lifting constraints on a person's capability set. A norm or rule can lead either to a limitation of a person's functionings, such as in case of a social norm that women are not allowed to ride a bicycle, or to an expansion of a person's functionings, such as in case of rules for mandatory vaccination by which a person's capability set is expanded to make the functioning "living in a measles-free world" accessible. Rules or norms are thus not by definition freedom-limiting or freedom-expanding, but rather depend on the specific way they affect the access to certain functionings.

This aspect allows us, for instance, to assess certain rules imposed on economic systems in a different light. Consider the debate about free trade: while some extreme positions would consider any (global or national) regulations imposed on trade or financial transactions as a risk to the creation of wealth and development, others argue the opposite and defend a reversal from free trade or globalization to a protectionist economic policy to enhance the well-being of people of the respective societies. The capability approach would consider which rules contribute to the expansion of human capability (such as joint rules on minimal social and environmental standards or corporate taxation avoiding a race to the bottom) and which ones risk to limit the expansion of human capability and development (as it might be the case with a certain class of protective tariffs).

A fourth characterizing feature of the capability approach that makes an important difference in the assessment of economic systems is its focus on ends instead of means. The main argument in the capability literature that motivates this shift from means towards ends is the need to account for human diversity in the conversion of means into ends (Sen 1992). However, in the assessment of economic systems this shift can have another merit as well, namely to ask how our institutional structures (or other factors that influence conversion factors, such as social or legal norms) can influence the conversion of material resources (means) into functionings (ends). This question has particularly high relevance in the light of climate change and other effects of our current economic system on the environment and resource use. The capability framework, with its distinction between resources and well-being, allows one to shed light on how our social, economic, and legal institutions affect the amount of resources we need to reach a certain level of well-being. It allows one to investigate whether we can change our institutional structures and norms in such a way that decreases the resources we need to obtain a certain level of well-being, that is human functionings. The resource use necessary

to realize the functioning “moving around,” for instance, will be very different in a society that is mainly dependent on private car use than in one that has a very good and affordable public transport network. This does not mean that the decrease in resource use needed to prevent irreversible environmental damage can be achieved without any decrease of human well-being as conceived by many. However, it does mean that the decrease in human well-being is not as large as it would be if one were to measure human well-being purely by using a proxy such as the resources available to a person (or generated in a society by means of GDP growth) without accounting for factors that affect the conversion of these means into the ultimate ends of well-being.

So far, we discussed how some of the characterizing features of the capability approach—namely the broad notions of constraints it accounts for in its conception of freedom, its general approach to conversion factors in their possible capability-expanding or capability-constraining impact, and the focus on ends instead of means—make a difference in the analysis of economic institutions and systems. Of course, the outcome of an assessment of economic systems in terms of the capability approach will crucially depend on how the framework is further specified and on the precise question it is meant to address. Instead of speculating about outcomes of such specific assessments, an exercise which would exceed the scope of this chapter and is left to future research, we shall move next to a number of unresolved questions in the specification of the capability approach necessary for its use in the assessment of economic systems, and point to some limitations of using the capability approach in economic ethics in the concluding section.

5.5.2 How Should the Capability Approach Be Further Developed for the Assessment of Economic Institutions and Systems?

Let us now turn to the specific exercise of assessing economic systems, as discussed in the previous section. What are the open questions that remain to be answered in order to sufficiently specify the capabilitarian framework and put the capability approach to use in the assessment of economic systems and institutions more generally? The recent literature in welfare economics that aims to expand the informational basis of welfare assessments in line with the capability approach is very useful to clarify these open questions and make them more precise. The assessment of societal mechanisms and systems is a crucial aspect of welfare economics. Indeed, contrary to widely held perceptions, the reach of welfare economics extends far beyond the assessment of the market mechanism. In this light it should come as no surprise that the capability approach has one of its roots in welfare economics, namely as an answer on how to broaden the informational basis of welfare judgements, moving beyond the common framework of so-called welfarism, much criticized by Sen. Welfarism demands the evaluation of social states in

terms of a person's preferences. Sen criticized this informational poverty of welfare economics and argued for enlarging the informational space by considering information about people's rights and freedoms in one's assessment as well (Sen 1970). The capability approach can be seen as one way the informational basis can be extended.

One of the merits of the formal framework that is employed in welfare economics is that it puts unresolved questions or inconsistencies that are often difficult to detect in a verbal argument out into the open. Without going into great depth of the formal framework of this literature (see Dowding and van Hees 2009), we want to briefly discuss some of the unresolved questions identified there (Pattanaik and Xu, forthcoming) and elsewhere (Basu 1987) that need to be answered before economic systems can be systematically assessed (formally and informally) with the capability approach.

One crucial unanswered question is how to account for the interpersonal dependency of capability sets. If the availability of a functioning in a person's capability set is dependent on the actions of others, how can it be identified without knowing the choice of other people from their capability sets? For example, it often depends on the action of a person's partner whether the functioning to make career and have children is indeed available to a person, because in the current institutional setting, often it is not available to both partners.

Another open question is how interpersonal comparisons of capability allocations can or should be made. If a certain policy leads to the expansion of the capability sets of one group at the cost of a limitation of capability sets of another group, how can or should they be compared (and possibly) traded off against each other? There are different solutions to this problem. One is to consider only Pareto-wise capability improvements—that is, policies that expand the capability sets of some without diminishing the capability set of anybody else in society. Beside the fact that this would lead to very incomplete rankings of economic systems and policies in terms of capability, it would also suffer a well-known problem of any Pareto criterion, namely that it does not account for the distribution of capability among different groups in society.

A different set of challenges in order to further develop the capability approach for the evaluation of economic institutions and systems lies on the empirical front. The capability approach is theoretically and normatively attractive because of its informational riches, but this same fact also makes it very challenging on the empirical front (Kuklys 2005). For many functionings that we have good reasons to believe are relevant when evaluating economic institutions and systems, we either do not have very good ways to measure it, or else these empirical measures require much more radical and intensive multidisciplinary collaboration than is currently operating. For example, not suffering from avoidable and harmful levels of stress seems a very relevant desideratum of economic systems, but it is not something that economists generally know how to measure. But even if the incentives that hamper far-reaching multidisciplinary collaboration in the measurement of well-being and effective freedom were effectively addressed, there will remain additional challenges to measuring the impacts on people's functionings of a change in an economic institution. For example, there are limits to how long surveys gathering data can be if one does not want to

lose too many respondents. This is a challenge to the capability approach, because a complete capability assessment will require a lot of information about the effects of a policy or institutional change on the affected persons. One can wonder whether this is possible at an affordable cost, and whether such a long survey will negatively affect its response rate.

In this section we have not aimed at giving a complete overview of all the questions that remain to be answered before the full potential of the capability approach for economic ethics can be harvested. However, the selection of questions that we discussed already gives a clear indication that much work remains to be done.

5.6 CONCLUSION

The description and analysis of the capability approach as we presented it in this chapter provides reasons to believe that it is well-suited for questions in economic ethics in which well-being and effective freedom play an important role. But one should be careful in the conclusions that one draws regarding the potential reach or scope of the capability approach for the study of questions in economic ethics.

For some questions in economic ethics, most of the normative analysis is done by other values or moral principles. Take the question of whether there should be a market for surrogates' services. In order to provide a sound ethical analysis of this question, one would have to engage with issues of bodily integrity, the meaning of life, what it means to be a parent and a child, parent–child relationships, the limits of commodification for bodily services, autonomy and exploitation, and so forth. Using some standard terminology from philosophical ethics, these are questions where non-consequentialist moral principles, as well as questions about “the right” rather than about “the good,” need to be addressed. As we have argued elsewhere, the capability approach focuses on questions about the good, but has nothing specific to offer to address questions that, within moral philosophy, typically fall under the heading of the right (Robeyns 2017: 143–5). It is therefore important not to overstretch the reach of the capability approach: it will not be a sufficient framework to address issues of economic ethics that fall into that category, although it does have an account of well-being to offer when addressing those questions (Satz 2010).

Despite this caveat, we believe that there are strong reasons to pursue the research agenda that we sketched in Section 5.5.2, and to develop systematic and careful answers to the open questions that remain to be answered before the capability approach can be systematically used in the assessment of economic systems and institutions in economic ethics. These open questions, which await resolution, constitute fascinating ground for new research in economic ethics and the capability literature. What exactly the limitations are of the capabilitarian framework in reference to economic ethics in general, and the assessment of economic systems and institutions in particular, can only be discovered by pursuing these questions.

REFERENCES

- Anderson, Elizabeth. 2000. "Optional Freedoms." In Joshua Cohen and Joel Rogers (eds.), *What's Wrong with a Free Lunch?* (Boston, MA: Beacon Press), pp. 70–4.
- Angner, Erik. 2015. "Well-Being and Economics." In Guy Fletcher (ed.), *The Routledge Handbook of the Philosophy of Well-Being* (London: Routledge), pp. 492–503.
- Arneson, Richard. 2010. "Two Cheers for Capabilities." In Harry Brighouse and Ingrid Robeyns (eds.), *Measuring Justice: Primary Goods and Capabilities* (Cambridge: Cambridge University Press), pp. 101–27.
- Basu, Kaushik. 1987. "Achievements, Capabilities and the Concept of Well-Being." *Social Choice and Welfare* 4: 69–76.
- Berlin, Isaiah. 1969. *Four Essays on Liberty*. Oxford: Oxford University Press.
- Binder, Constanze. 2009. "Context Dependency of Valuable Functionings: How Culture Affects the Capability Framework." In Enrica Chiappero-Martinetti (ed.), *Debating Global Society: Reach and Limits of the Capability Approach* (Milan: Giangiacomo Feltrinelli Editore), pp. 203–29.
- Binder, Christina, and Constanze Binder. 2017. "A Capability Perspective on Indigenous Autonomy." *Oxford Development Studies* 44 (3): 1–18.
- Byskov, Morten F. 2017. "Democracy, Philosophy, and the Selection of Capabilities." *Journal of Human Development and Capabilities* 18: 1–16.
- Carter, Ian. 1999. *A Measure of Freedom*. Cambridge: Cambridge University Press.
- Claassen, Rutger. 2009. "Institutional Pluralism and the Limits of the Market." *Politics, Philosophy & Economics* 8: 420–47.
- Claassen, Rutger. 2017. "An Agency-Based Capability Theory of Justice." *European Journal of Philosophy* 25: 1279–304.
- Cohen, Gerald A. 1981. "Freedom, Justice, and Capitalism." *New Left Review* 126: 3–16.
- Cohen, Gerald A. 1989. "On the Currency of Egalitarian Justice." *Ethics* 99: 906–44.
- Crocker, David A. 2008. *Ethics of Global Development: Agency, Capability and Participatory Democracy*. New York: Cambridge University Press.
- Dowding, Keith, and Martin van Hees. 2009. "Freedom of Choice." In Paul Anand, Prasanta K. Pattanaik, and Clemens Puppe (eds.), *The Handbook of Rational and Social Choice* (Oxford: Oxford University Press), pp. 374–92.
- Folbre, Nancy. 2001. *The Invisible Heart. Economics and Family Values*. New York: New Press.
- Hayek, F. A. 1960. *The Constitution of Liberty*. Chicago, IL: University of Chicago Press.
- Ibrahim, Solava. 2017. "How to Build Collective Capabilities: The 3C-Model for Grassroots-Led Development." *Journal of Human Development and Capabilities* 18: 197–222.
- Kuklys, Wiebke. 2005. *Amartya Sen's Capability Approach. Theoretical Insights and Empirical Applications*. Berlin: Springer.
- MacCallum, Gerald G., Jr. 1967. "Negative and Positive Freedom." *Philosophical Review* 76: 312–34.
- Nozick, Robert. 1974. *Anarchy, State and Utopia*. New York: Basic Books.
- Nussbaum, Martha. 2000. *Women and Human Development*. Cambridge: Cambridge University Press.
- Nussbaum, Martha. 2006. *Frontiers of Justice*. Cambridge, MA: Belknap Press of Harvard University Press.
- Nussbaum, Martha. 2011. *Creating Capabilities: The Human Development Approach*. Cambridge, MA: Belknap Press of Harvard University Press.

- Oppenheim, Felix. 1985. "Constraints on Freedom' as a Descriptive Concept." *Ethics* 95: 305–9.
- Parfit, Derek. 1984. *Reasons and Persons*. Oxford: Oxford University Press.
- Pattanaik, Prasanta K., and Yongsheng Xu (forthcoming). "On Capability and Its Measurement." In E. Chiappero-Martinetti, S. Osmani, and M. Qizilbash (eds.), *The Cambridge Handbook of the Capability Approach* (Cambridge: Cambridge University Press), in press.
- Qizilbash, Mozaffar. 2013. "On Capability and the Good Life: Theoretical Debates and Their Practical Implications." *Philosophy and Public Policy Quarterly* 31(2): 35–42.
- Robeyns, Ingrid. 2016a. "Capabilitarianism." *Journal of Human Development and Capabilities* 17: 397–414.
- Robeyns, Ingrid. 2016b. "The Capability Approach." In Edward N. Zalta (ed.), *Stanford Encyclopedia of Philosophy*. Available at <https://plato.stanford.edu/archives/win2016/entries/capability-approach/>.
- Robeyns, Ingrid. 2017. *Well-Being, Freedom and Social Justice: The Capability Approach Re-Examined*. Cambridge: Open Book Publishers.
- Satz, Debra. 2010. *Why Some Things Should Not Be for Sale: The Moral Limits of the Market*. New York: Oxford University Press.
- Sen, Amartya. 1970. *Collective Choice and Social Welfare*. London: Penguin Books (2017 edition).
- Sen, Amartya. 1980. "Equality of What?" In Sterling M. McMurrin (ed.), *The Tanner Lectures on Human Values*, pp. 196–220. Salt Lake City, UT: University of Utah Press.
- Sen, Amartya. 1984. "The Living Standard." *Oxford Economic Papers* 36: 74–90.
- Sen, Amartya. 1985a. *Commodities and Capabilities*. Oxford: Oxford University Press.
- Sen, Amartya. 1985b. "Well-Being, Agency and Freedom: The Dewey Lectures 1984." *Journal of Philosophy* 82: 169–221.
- Sen, Amartya. 1992. *Inequality Reexamined*. Cambridge, MA: Harvard University Press.
- Sen, Amartya. 1993a. "Capability and Well-Being." In Martha Nussbaum and Amartya Sen (eds.), *The Quality of Life* (Oxford: Clarendon Press), pp. 30–53.
- Sen, Amartya. 1993b. "Markets and Freedoms: Achievements and Limitations of the Market Mechanism in Promoting Individual Freedoms." *Oxford Economic Papers* 45: 519–41.
- Sen, Amartya. 2002. *Rationality and Freedom*. Cambridge, MA: Harvard University Press.
- Sen, Amartya. 2004. "Capabilities, Lists, and Public Reason: Continuing the Conversation." *Feminist Economics* 10(3): 77–80.
- Sen, Amartya. 2009. *The Idea of Justice*. Cambridge, MA: Harvard University Press.
- Sen, Amartya, and Sunil Sengupta. 1983. "Malnutrition of Rural Children and the Sex Bias." *Economic and Political Weekly* 18: 855–64.
- Van Zanden, Jan Luiten, Joerg Baten, Marco Mira d'Ercole, Auke Rijpma, Conal Smith, and Marcel Timmer. 2014. *How Was Life? Global Well-Being Since 1820*. Paris: Organisation for Economic Co-operation and Development.

B. SOURCES OF MORALITY IN ECONOMICS

CHAPTER 6

EVOLUTION AND MORAL MOTIVATION IN ECONOMICS

GEOFFREY M. HODGSON

6.1 INTRODUCTION

IN both his *Theory of Moral Sentiments* (1759) and *Wealth of Nations* (1776), Adam Smith rejected the idea that humans were motivated by pleasure and self-interest alone. Although we often act selfishly, we have guiding “moral sentiments” and “sympathy” as well. But the evolution of economics since Smith has seen the rise of the narrower view of the individual as a utility-maximizer. Morality is widely subsumed under individual preferences or utility.¹ Furthermore, the specific version of utility maximization that until recently has dominated economics is one of the self-regarding or self-interested individual, who gains satisfaction simply from his or her own advantage or gain. A selfish and non-altruistic version of utility maximization has prevailed.

Although there were several predecessors in the history of economic thought, this shift in thinking was announced in two seminal works by William Stanley Jevons (1871) and Carl Menger (1871), who placed individual self-interest at the foundation of economics. Three years later, Léon Walras (1874) built neoclassical general equilibrium analysis upon a similar assumption of self-interest. For the next 100 years or more, self-interested “economic man” was the centrepiece of mainstream economic theory.

In the same pivotal year, however, Charles Darwin (1871) published a contrasting and evolutionary explanation of cooperative solidarity and morality, which took over 100 years to be confirmed broadly by theoretical and empirical research. Interestingly, Darwin (1871: vol. 1, 81) was partly inspired by Smith’s account of morality in the *Theory of Moral Sentiments*.

¹ One of the few prominent dissenters to this still-dominant utilitarian view is Sen (1977, 1987).

Many economists now accept that this narrow, self-regarding view of humans is flawed. There is now an enormous body of empirical research confirming that humans have cooperative as well as self-interested dispositions.² Altruism is often explained in terms of an “other-regarding” preference function, where an individual gains utility from observing the satisfaction of others. Cooperation can occur out of self-interest or altruism.

But many accounts conflate morality with altruism or cooperation. Morality is still treated as an aspect of altruistic or cooperative behavior in a utility-maximizing framework.³ Yet Smith and Darwin argued that morality was an additional factor, irreducible to utility and preference. We need to understand why people invoke morality and follow moral rules for reasons other than preference or self-interest.

Morality remains controversial for moral philosophers, notwithstanding a widespread view among them that moral judgments cannot be treated as matters of mere preference or utility maximization. Morality means “doing the right thing.” It entails notions of justice that can trump our preferences or interests. Moral judgments are by their nature inescapable, and would apply to all in the same circumstances. They are buttressed by both emotional feelings and reasoned argument. Consequently, morality differs fundamentally from matters of mere convenience, convention, or conformism. Moral feelings are enhanced by learned cultural norms and rules. Morality is a group phenomenon involving deliberative, emotionally driven and purportedly inescapable rules that apply to a community.

A society or economy cannot function without moral bonds and rules. Our understanding of social institutions and organizations is inadequate unless we appreciate the moral motivations of individuals within them, and how those institutions help to sustain and replicate these moral sentiments.

Accepting the importance of morality does not mean denying that individuals are self-interested. We are selfish, to a large extent. But we are also moral beings, and our ethical feelings and beliefs, well-formed or otherwise, play an ubiquitous role in our interactions with others, even in the modern acquisitive world of business and consumerism.

It is necessary to bring moral motivations back into the picture. Fortunately there is now a flourishing stream of recent evolutionary studies—spanning several disciplines including anthropology, primatology, philosophy, and economics—that can help to fill the gaps. We can begin to understand the role of morality in sustaining institutions and enhancing social cohesion in any society.

In recent years our understanding of the evolution of cooperation has developed enormously. There are many powerful models and simulations exploring mechanisms and conditions under which altruistic or cooperative individuals can become established in a population. What does a focus on morality add to this? By assuming that behavior

² See, for example, Güth (1995), Field (2001, 2007), Henrich et al. (2001, 2004), Hammerstein (2003), and Bowles and Gintis (2011).

³ See, for example, Fehr and Gächter (2002), Bowles et al. (2003), Boyd et al. (2003), Gintis et al. (2005), Kaplow and Shavell (2007), Bowles and Gintis (2011), and Gintis and Helbing (2015b).

largely reflects inherited dispositions and neglecting the possibility of countervailing moral deliberations, prominent models of the evolution of altruism pay insufficient attention to why an intelligent and thoughtful individual should bear costs to help others, even if she was genetically and emotionally disposed to do so. Although the modelers of cooperation often mention morality, they fail to illuminate its crucial role, typically by forcing matters of ethics and preference together into one utilitarian box.⁴

When we turn our attention from the past evolution of morality in humans to the consideration of economic and social policies for the present, then a deeper and more detailed understanding of human motivation is required. By treating motivation as all stemming from individual “preferences”—as is typical in much of the literature on the evolution of human cooperation—approaches to policy are more easily diverted into the narrower channels of material or pecuniary incentives, neglecting moral motivations and appeals to ethical values.

The nature of morality is discussed in the following section, while the evolution of morality is considered in the following three sections. The narrative moves from Darwin’s sophisticated view to the modern proposition that moral feelings have a genetic foundation as well as a heavy dependence on culture and deliberation. A final section concludes the chapter.⁵

6.2 WHAT IS MORALITY?

Morality is complex and controversial. In Darwin’s (1871: vol. 1, 87–9) account, morality results from a combination of emotional impulses and thoughtful deliberation. He argues that although primitive moral feelings have evolved for millions of years among “the progenitors of man,” humans alone have a developed sense of morality:

A moral being is one who is capable of comparing his past and future actions or motives, and of approving or disapproving of them. We have no reason to suppose that any of the lower animals have this capacity...man...alone can with certainty be ranked as a moral being. (162)

For Darwin, morality emerged in humans upon a long-evolved foundation of instinct and impulse. As I note in the following section, Darwin also saw morality as a social phenomenon, involving social relations and shared values.

⁴ See the references in the preceding endnote.

⁵ This chapter makes use of material from Hodgson (2013, 2014), which includes more extensive evidence and discussion. In Hodgson (2013) I also contrast my own views on morality with others. For example, in leading accounts by Etzioni (1988) and Sen (1987) there is no attempt to explain how our moral or other dispositions evolved. By contrast, Binmore (1994, 1998) develops a sophisticated evolutionary theory but reduces morality to matters of mere convention. I thank many colleagues, too numerous to mention, for helpful discussion and feedback on these issues.

Much of the recent theoretical work by economists that attempts to explain cooperation in the real world conflates issues of morality with altruism or cooperation under the description of “social” or “other-regarding” preferences (Bowles and Gintis 2011; Camerer and Fehr 2006; Fehr and Fischbacher 2002; Fehr and Camerer 2007). The assumption of other-regarding preferences contrasts with the previously prominent idea that economic man was entirely selfish. But someone with other-regarding preferences is still maximizing *her own* utility, and may be regarded as selfish too.

To some extent the work of Bruno Frey (1997) improves on this with his distinction between extrinsic and intrinsic motivation. Morality relates more closely to the latter than the former. Viktor Vanberg shows that once we distinguish between preferences over outcomes and preferences over actions then morally motivated behavior cannot fit into the former category:

There is...a significant difference between claiming, on the one hand, that agents evaluate outcomes not only in terms of their own narrowly defined interests but also in terms of how they affect the well-being of other persons, and claiming, on the other hand, that agents are motivated to act in accordance with ethical rules or principles of fairness. (2008: 608)

The incapacity of the standard utilitarian calculus—or even the language of preferences—to depict adequately the nature of morality becomes even more graphic if we consult the literature in ethical philosophy.

Many controversies divide moral philosophers. The best we can do here is to pick up rather selectively some relevant threads and some prominent descriptions of the nature of moral judgment. The leading moral philosopher Richard M. Hare (1952) argued that morality was subject to reason and one cannot hold contradictory ethical judgments. He also maintained that any normative judgment was universalizable in the context to which it pertained, in the sense that anyone proclaiming an “ought” in a particular context was committed to prescribing a similar normative judgment for anyone in any relevantly similar situation. As J. L. Mackie put it in his classic account, a moral judgment

is not purely descriptive, certainly not inert, but something that involves a call for action or for the refraining from action, and one that is absolute, not contingent upon any desire or preference or policy or choice, his own or anyone else’s. (1977: 33)

In his impressive philosophical account of the *Evolution of Morality*, Richard Joyce (2006: 70) argues on the basis of considerations in the philosophical literature that morality has most or all of the following characteristics:

1. Moral judgments express attitudes (such as approval or contempt) and also express beliefs.
2. The emotion of guilt is an important mechanism for regulating moral conduct.
3. Moral judgments transcend the interests or ends of those concerned.
4. Moral judgments imply notions of desert and justice.

5. Moral judgments are inescapable.
6. Moral judgments transcend human conventions.
7. Moral judgments govern interpersonal relations and counter self-regarding individualism.

These characteristics do not establish a *valid* morality; they instead help us to identify what is a *moral judgment*, whether acceptable or otherwise. The argument in this chapter relies on descriptive rather than normative ethics: there is no attempt here to identify the “right” morality, but instead to identify *the basic nature of a moral claim*. Most religions uphold moral claims, but that does not make them all right or just.

Like Darwin, Joyce emphasized the role of the emotions as well as deliberation. His first point establishes that a moral judgment must involve both beliefs and sentiments, and is not reducible to either alone. If an action is impelled *purely* by emotion and sentiment then—as Darwin understood—it cannot amount to moral motivation. Deliberations and beliefs are also vital, but are themselves insufficient because they must be backed by sentiments or emotions: acting morally is more than calculated conformity to moral rules.

Moral judgments may be rationalized in various ways, but they are more than matters of propositional belief or logical syllogism. Defiance of shared moral rules in a group is often met with emotional hostility. Conformity to them may sometimes bring a warm emotional glow. The emotional dimension of moral rules plays an important role in their evolution and their survival, as I shall discuss further below. Guilt (point 2) is a particularly important emotion that sometimes emerges after breaches of moral rules, and it too plays a part in the evolutionary process.

Joyce’s points 3 through 7 reveal the limitations of typical utilitarian approaches. Moral judgments are not simply expressions of an individual’s interests, preferences, sentiments or beliefs. They are also claims to universality in their context, which would apply irrespective of the interests, preferences, sentiments or beliefs of those to whom they are supposed to apply.

As both Mackie and Joyce insist, morality surpasses questions of preference. It is a matter of right or wrong, or of duty, of “doing the right thing,” irrespective of whether we like it or not. This is part of what makes us human: we are capable of considering moral rules, and understanding that their observance is more than a matter of personal whim or satisfaction. This dimension is missing in much of economics. Moral values are either ignored or subsumed under matters of utility or preference.

Modern society establishes a fundamental difference between moral rules and other (normative) rules. “Murder is wrong” does not carry the same connotations as “splitting infinitives is wrong” or “in Britain one must drive on the left side of the road.” Linguistic and traffic rules are matters of convention; they are non-universal. But punishment may still occur when some conventions are breached. Murder is also punishable, but by contrast it is more than a breach of convention.

Threat of punishment or respect for the law are each insufficient to explain the relatively low frequency of murder and other crimes. Most of us abstain from murder not simply because the probability of severe punishment outweighs any expected benefit.

Most of us refrain from murder because we believe that it is *morally wrong*; we would desist even if we lived in a country where murder went unpunished.

Although there is a difference between morality and mere convention, some conventional rules seem to acquire a moral imperative when they become laws. They sometimes inherit the force of morality from other purportedly universal moral rules, particularly the need to respect others and to obey the law. While conventions may differ from culture to culture, we often conform to them, partly out of mutual respect or legal responsibility. Hence matters of mere convention can acquire some moral force if they become enshrined in law. If so, they do not necessarily become moral issues themselves, but their observance may acquire moral substance by virtue of their legal status. Consequently the moral legitimacy (or otherwise) of the legal system in the eyes of citizens is crucial (Tyler 1990; Hodgson 2009, 2015a).

It is a commonplace observation that what may be a moral rule for one culture may not be so for another. But this does not mean that moral rules are reducible to conventions. They become moral rules because many people believe in them as such, and they jointly uphold them as more than matters of convenience, self-interest or convention. The cultural specificity of some moral judgments does not justify a *normative moral relativism*, where one person's morality is deemed as good as any other.⁶

Advocates of normative moral relativism take a further step and uphold that if an act is regarded as morally permissible in a culture then it must be deemed acceptable even if it is regarded as wrong elsewhere. To accept such a version of moral relativism is to undermine an essential feature of morality itself: that it is absolute and inescapable. Because this feature is denied, such a moral relativist cannot believe in *any* moral judgment in the terms defined above. Hence normative moral relativists are obliged to become moral nihilists or amoralists.

Just as there are unavoidable ontological commitments in any process of enquiry, for humans in societies there are unavoidable moral commitments. And some moral rules are ethically superior to others. Moral relativists respond that there is no way of knowing what the superior moral rules are. But even if there were no way of finding them, this argument is invalid. As with the epistemic error in ontology—the fact that we can never prove that the real world exists does not mean that there is not a real world outside our senses—there is a similar error in ethics—the fact that we may be unable to identify a superior moral code or prove that it is valid does not mean that a morally superior code does not exist. Ignorance of a valid morality does not mean that it is in principle unspecifiable.

Significantly from an evolutionary perspective, studies show a number of common features of moralities across cultures, notwithstanding other important cultural variations.⁷

⁶ The uncontroversial observation that different cultures have different ethical codes can be described as *descriptive moral relativism*. See Hodgson (2013) for a discussion.

⁷ Bok (1978), Roberts (1979), Brown (1991), Schwartz (1994), Haidt and Joseph (2004), Nichols (2004), and Haidt (2012).

All cultures regard many acts of harm against others as immoral and invest many acts of reciprocity and fairness with moral virtue. All cultures have moral rules concerning required behaviors specific to particular social positions, roles, or ranks. Moral codes restraining individual selfishness are also commonplace. As well as sustaining enormous cultural diversity, genetic and cultural co-evolution has ensured that some specific types of pro-social moral rules have endured.

In summary, and in answer to the question that heads this section, a moral judgment involves an expression of attitudes, beliefs, and emotions but is also subject to deliberation concerning matters of fairness or justice. In contrast to standard utilitarian approaches, a moral judgment is more than mere convention; it is inescapable and transcends individual preferences or interests.⁸ A moral system refers to shared and interactively reinforced moral values in a society or social group. These definitions are incomplete and imprecise, but they are sufficient for our purposes in this chapter.

6.3 DARWIN AND THE EVOLUTION OF MORALITY

We need evolutionary explanations of the origin and persistence of morality. Darwin's account of this evolution is remarkably resilient, even in the light of modern research. In his *Descent of Man*, Darwin (1871: vol. 1, 162) considered dispositions such as "sympathy, fidelity, and courage" that would advantage one tribe against the other in their struggle for existence, and which had been originally "acquired by the progenitors of man." The accent on sympathy is of course reminiscent of Adam Smith in the *Theory of Moral Sentiments*.⁹ Among the other qualities he considers, Darwin listed the disposition to obey those in authority rather than follow individually selfish motives:

Obedience . . . is of the highest value, for any form of government is better than none. Selfish and contentious people will not cohere, and without coherence nothing can be effected. A tribe possessing the above qualities in a high degree would spread and be victorious over other tribes. (1871: vol. 1, 162)

⁸ Utilitarians will not accept this. But I cannot develop a full critique of utilitarianism here. The behavior of any real entity can be made consistent with some utility function (including, more controversially, apparently "inconsistent" behavior). For utilitarians that is enough to clinch the matter. But for the critics, fitting a function to behavioral data does not explain anything, and it overlooks that which specifically makes us human. For fuller discussion and further references see Smart and Williams (1973), Sen (1987), and Hodgson (2013).

⁹ In his *Moral Sentiments* and elsewhere Smith emphasized moral motivations and the importance of justice in economic arrangements (Sen 1987; Evensky 2005). Darwin made notes on the *Moral Sentiments* but there is no evidence he read the *Wealth of Nations*. He famously read Malthus's *Essay on the Principle of Population*, but the pages of the copy of the first volume of *Capital* sent to him personally by Marx remained uncut (Hodgson 1993).

Obedience to authority is of course a vital mechanism in the establishment of a system of morality in society. Darwin wrote:

It must not be forgotten that although a high standard of morality gives but a slight or no advantage to each individual man and his children over the other men of the same tribe, yet that an advancement in the standard of morality and an increase in the number of well-endowed men will certainly give an immense advantage to one tribe over another. There can be no doubt that a tribe including many members who, from possessing in a high degree the spirit of patriotism, fidelity, obedience, courage, and sympathy, were always ready to give aid to each other and to sacrifice themselves for the common good, would be victorious over most other tribes; and this would be natural selection. (1871: vol. 1, 166)

Darwin proposed that groups containing individuals that devote themselves to the interests of their group will have an advantage in the struggle for survival. Among humans, binding sentiments of sympathy and solidarity are strengthened by a moral code, typically transmitted by instruction and often sanctified by religion (Wilson 2002).

Darwin's evolutionary explanation of moral sentiments relies to some degree on a notion of group selection, where individual traits that benefit the group are assumed to prosper. Darwin did not counter the objection that selfish individuals would be able to free-ride within an altruistic group, and eventually out-breed the unselfish (Williams 1966; Dawkins 1976). Consequently, before the theory of group selection was rehabilitated, Darwin's theory of the evolution of morality was regarded as quaint and outmoded.¹⁰ Darwin's account also suffered because of a long-established and popular contrary view that the foundations of human morality are of recent origin, rather than based on "social qualities . . . acquired by the progenitors of man."

Darwin was ignorant of the mechanisms of inheritance, including genes. We now understand that genetic group selection must be distinguished from cultural group selection. As Joseph Henrich (2004) shows, both types of group selection are possible in principle under specified conditions, but this does not always necessarily mean that they are always strong. Furthermore, genes and culture interact with one another in specific ways. Culture provides part of the environment in which genes are selected, and our genetic endowment influences cultural evolution (Boyd and Richerson 1985; Durham 1991).

Strong arguments support the notion of cultural group selection among humans (Boyd and Richerson 1985; Henrich 2004). The existence of genetic group selection in our species is more problematic. Although it is possible in principle, it depends on the restriction of inter-group migration and the limitation of genetic mixing between groups. But the evidence among primates is that significant group-to-group migration does occur (De Waal 2006: 16). There is also inter-group migration in some contemporary hunter-gatherer societies, and consequently lower levels of genetic relatedness within groups (Hill et al. 2011). We lack any clear evidence on the degree of inter-group

¹⁰ Major contributions to this rehabilitation include Wade (1978), Wilson (1980, 1983), Wilson and Sober (1994), Sober and Wilson (1998), and Wilson and Wilson (2007).

migration among early humans, but we have no reason to presume that they differed radically from primates in this respect, although several hypotheses have been developed (Bowles and Gintis 2011).

Consequently, on the basis of existing evidence, the genetic foundations of altruistic and moral feelings seem more likely to have evolved first through mechanisms of kin altruism and then reciprocal altruism. William Hamilton (1964) demonstrated that altruistic genetic dispositions can evolve among closely related individuals and Robert Trivers (1971) showed how altruism could be reinforced in small groups by mechanisms sustaining reciprocation. Our best guess with current knowledge is that altruistic, cooperative, and moral feelings then required the further emergence of a culture, so that they could spread through the group and become reinforced by enduring cultural norms. In short, genetic mechanisms established critical masses of altruists in social groups, leading to the spread of cultural norms sustaining cooperation and to the development of systems of morality that further enhanced the fitness of groups. Genes played a role, but also indispensable was culture, particularly through the inculcation of behavioral norms in children by parents (Palmer and Steadman 1997).

In any case it seems certain that very basic moral feelings have a genetic basis and have evolved in family and kin groups. Because reciprocity and cooperation in such circumstances enhances the fitness of the genes, emotional and other dispositions that aided cooperation and family cohesion also had a survival advantage. Moral sentiments thus evolved on a genetic foundation. But they require structured social interaction to become channeled and expressed. Culture developed these sentiments into a transmitted moral code. Hence the evolutionary origins of morality involved the interplay of genetic and cultural factors. Morality thus has both individual and social aspects. In the following two sections the genetic foundation of morality is discussed in more depth.

6.4 CONTRA HUXLEY AND DAWKINS: MORALITY IS MORE THAN SKIN DEEP

Is morality something that emerged only after the development of complex societies and civilized behavior? Is it a recent innovation that has to be instilled into everyone against their rude and selfish evolved nature? Or is it more than skin deep?

Darwin insisted that morality proper relies on levels of conscious deliberation, which are developed among humans alone. But at the same time, he saw the biological roots of morality as a product of natural selection. Because of our unique capacities for linguistic communication, humans were the first species to develop and articulate moral codes; but the foundations of morality go far back into our pre-human past.

However, evolution and morality have often been regarded as separate issues, even by Darwin's disciples. In 1893, Darwin's friend Thomas Henry Huxley famously argued that "the ethical progress of society" depends on neither ignoring, accepting nor imitating

natural selection in that sphere “but in combating it” (Huxley and Huxley 1947: 82).¹¹ Redolent of Huxley, Richard Dawkins (1976: 2–4, 215) describes the “ruthless selfishness” of our genes. He also claims that “we are born selfish” and “anything that has evolved by natural selection should be selfish.” Not only genes but also individuals are described as selfish.¹² Similar claims that evolution gives rise to the self-interested human individual can be found in works by Michael Ghiselin (1974) and Richard Alexander (1987). But Dawkins adds in a contradictory and Huxleyan manner: “Let us try to teach generosity and altruism. . . . We, alone on earth, can rebel against the tyranny of the selfish replicators.” We are born selfish but somehow we can choose to be otherwise. The contrast with Darwin’s (1871) views on human nature and morality are graphic.

There are many problems with the Huxley-Dawkins separation of our evolutionary legacy from our current choices and capacities (Midgley 2003; De Waal 2006; Joyce 2006). Above all, if natural selection provides us with selfish dispositions, then why should we be inclined to “combat” selfishness or “teach generosity and altruism,” unless they are some sort of elaborate selfish ruse to get the upper hand? Why should nature-red-in-tooth-and-claw rule in one sphere but not another? Furthermore, any inclinations “to teach generosity and altruism” and to “rebel” against our own selfishness are unexplained. We are asked to overturn natural selection in the human domain without the assistance of any evolved human attribute. To “combat” natural selection would also mean suppressing the evolved moral sense and the only possible moral ally in this rebellion against selfishness.

Frans De Waal (2006: 10) criticizes the Huxley-Dawkins view where “morality is presented as a thin crust underneath which boil antisocial, amoral and egotistic passions.” He describes this as the “veneer theory” of morality and presents contrary evidence from his studies of primates that they are capable of sympathetic and cooperative emotions. Hence our moral capacities are grounded in our evolution as a social species, over many millions of years.

On one side is the “veneer theory” view that evolution leads to individual selfishness only, and morality is a desirable but optional overlay. On the other side, the critics argue that our long survival as a social species has meant the evolution of social instincts such as fairness and sympathy and the emotional capacity for guilt, and these form part of the foundations for morality. The critics cite evidence from the study of primate behavior from psychology, neuroscience, and experimental economics.¹³

¹¹ Huxley’s severance of the evolution of ethics from the evolution of humanity was countered by a number of writers including Westermarck (1891) and Kropotkin (1902).

¹² This is a major inconsistency in Dawkins’ (1976) position. The “selfish gene” metaphor supports his notion that evolutionary selection prioritizes successful genes because these are the key bits of information that replicate and program the growth and behavior of organisms. Whatever the merits and demerits of this “genes-eye view” of evolution, it cannot sustain the notion that individuals are entirely selfish as well. Inconsistently with his selfishness claim, Dawkins cites and endorses key arguments by Hamilton (1964) and Trivers (1971) that show that the individual may bear costs or even sacrifice itself, as if it were acting instead for the sake of its genes.

¹³ See, for example, Güth (1995), De Waal (1996), Greene and Haidt (2002), Zak (2004), Tancredi (2005), Hauser (2006), and Fehr and Camerer (2007).

The evidence of the critics is compelling. But we may not conclude that a fully developed morality has existed for millions of years. Morality proper requires the capacity to compare and deliberate upon actions and motives. It requires communication with others and a highly developed language (Joyce 2006). As Darwin (1871: vol. 1, 72) wrote: “after the power of language had been acquired and the wishes of the same community could be distinctly expressed, the common opinion how each member ought to act for the public good, would naturally become to a large extent the guide to action.”

Morality proper, involving articulated abstract rules, could not have developed earlier than language. A sufficiently well-developed language is of fairly recent origin, the earliest estimates being hundreds of thousands of years (Oppenheimer 2004) and the more recent being around 50,000 years ago (Diamond 1991). By contrast, the genus *homo* has been in existence for several million years. So is morality a recently acquired “veneer” after all? To answer this question we need to distinguish between, on the one hand, the evolution of the preconditions for the development of moral systems and, on the other hand, the full emergence of morality itself. As I outline in the next section, some of the preconditions of our morality evolved as early as our ape-like ancestors.

6.5 HOW MORALITY EVOLVED: GENES AND CULTURE

The nuts and bolts of morality include capacities for sympathy and cooperation that we share with the primates. The evidence suggests that much of this is biologically inherited as social instincts. But long before the evolution of human language, a form of proto-cultural and non-genetic transmission occurs. De Waal (1996, 2006) argues that primates can read or even transmit emotional states such as approval, empathy and fear, through sounds, body language, facial expressions, and pheromonal excretions. This is known as “emotional contagion.” The communication of emotional states is at the core of the capacity for empathy: apes can understand and even share the joys or sufferings of others.¹⁴

The possibility of reading emotions in others, and even replicating emotional states, is a crucial inheritance mechanism for transmitting useful information and enhancing social cohesion. For example, the transmission of fear in the group can lead to collective flight from a predator. Such mechanisms are ubiquitous in nature. The capacities to respond to such signals are genetically inherited as instincts. Organisms are genetically programmed to respond to signals that are relevant for their survival.

Among the repertoires of response for sophisticated organisms is the imitation of the behavior of others. There are at least two types of imitation among humans and primates. The first is conformist transmission (Boyd and Richerson 1985). It has been shown that

¹⁴ Consider also *neural mirroring*, where emotional contagion is shown to be a result of “mirror neurons” in the brain (Rizzolatti et al. 2007).

genes disposing individuals to such conformism would be selected in some contexts (Henrich and Boyd 1998). A second psychological mechanism is prestige-based imitation (Henrich and Gil-White 2001; Henrich 2004), where individuals learn advantageously from the more successful. This second mechanism must involve capabilities to recognize social hierarchy and prestige. In any social species such instinctive propensities are likely to be selected over time; they would bestow survival advantages for the individual and the group.

There is also evidence for learned or inherited dispositions, which are triggered in specific contexts, to punish those who break the rules or fail to enforce them.¹⁵ The relevant inherited dispositions have evolved in our social species over millions of years. Some punishment involves “strong reciprocity” (Gintis 2000) where there is a propensity not only to punish cheats, free-riders, rule-breakers and self-aggrandizers, but also to punish others who fail to punish the offenders. Especially within small groups, these propensities are driven by strong emotional feelings of anger. There is also evidence of such dispositions among primates (De Waal 1996). In a complex culture, emotionally empowered rules can help to enhance notions of justice and morality (Darwin 1871; De Waal 2006; Robinson et al. 2007). Given these emotions were present in our ape-like ancestors, morality has a genetic as well as a cultural foundation. Genes are insufficient to generate a moral system, but the cultural phenomenon of morality is fueled by biologically grounded emotions and value impulses.

Support for this view comes from experiments showing that preverbal babies assess individuals in the light of their behavior towards others. Studies of six- and ten-month-old infants show that they prefer an individual who helps over one who hinders another, prefer a helping individual to a neutral individual, and prefer a neutral to a hindering individual (Hamlin et al. 2007). Such inclinations may serve as a foundation for later moral thought and action, and their early developmental emergence supports the view that such capacities have an inherited biological grounding. Other studies show that infants as young as four years have a sense of fairness (McCrink et al. 2010). Young children are capable of distinguishing moral from prudential norms (Tisak and Turiel 1984), although many adults would have difficulty articulating the distinction.

Moral development starts in very young children and is subject to abrupt developmental leaps. For example, ideas concerning fairness tend to emerge around the age of four years, to be applied wildly to multiple contexts (Fiske 1991). As Jonathan Haidt (2001: 827) concludes: “This pattern of sudden similarly timed emergence with overgeneralization suggests the maturation of an endogenous ability rather than the learning of a set of cultural norms.”

Through our genes we inherit the capacity to quickly respond to social dilemmas by developing emotions. These emotions not only dispose us to make choices but help us to

¹⁵ See Boyd and Richerson (1992), Andreoni (1995), De Waal (1996), Ben-Ner and Putterman (2000), Fehr and Gächter (2000a, 2000b, 2002), Gintis (2000), Field (2001), Price et al. (2002), Boyd et al. (2003), Carpenter et al. (2004), Gintis et al. (2005), Wiessner (2005), Henrich et al. (2006), Fehr and Gintis (2007), Guzmán et al. (2007), Carpenter and Matthews (2009), Henrich et al. (2010), and Bowles and Gintis (2011).

form rapid judgments concerning what is morally right or wrong. Moral judgments help us justify our actions to others and to exhort others to approve or imitate. Genetic dispositions to deal with social dilemmas by developing emotionally charged value intuitions can thus have strong survival value.

Our genes do not tell us what is moral or immoral. We try to learn that through engagement in a social culture. The foundations of our moral capacity have evolved over the millions of years that we have been a social species. But in the last 100,000 years or so, the full development of human moral capacities has been highly dependent on particular cultural settings, allowing for multiple and contrasting moral systems on the basis of a common instinctive bedrock.

Given that we also inherit genetically a long-evolved capacity to imitate others and a capacity for empathy, we are likely to conform to strong moral claims, especially when made by high-status or numerous individuals. Cultural mechanisms lead to conformity within the group. Morality thus operates in a group context. Although rebellion against the prevailing moral rules is possible, clashes are more likely between groups with different cultures.

The evolution of moral capacities is much like the evolution of language, as well as depending on language itself. Because it would be impossible for an infant to learn all the rules of a language solely by reinforcement learning and cultural transmission, some very basic linguistic capacities must be inherited as instincts.¹⁶ Language acquisition depends on both inherited genetic capacities and interaction with others in specific cultural contexts. Similarly, the evolution of morality is “half-art and half-instinct.” When we judge an action as morally right or wrong, in part we rely on instincts and feelings. But this does mean that evolution creates a single morality for humankind. Variation between communities in their expressed moral norms is like variation between cultures in their spoken languages (Hauser 2006: 419–20).

Allen Buchanan and Russell Powell (2015) based much of their critique of evolutionary theories of morality on the argument the group selection forces would explain an intragroup but not a universal system of morality. Group selection cannot explain the evolution of moralities that apply to the human race as a whole. In response it must be pointed out that universal moral claims are either relatively recent historically, or were for a long while practically of lesser effect because of limited intercourse with other cultures. The struggle for a universal moral discourse is very much a modern project. It may be explained less by group selection and more by the cultural diffusion of moral claims to apply other peoples or cultures, who may previously have been classed as inferior and beyond the moral code. Such a diffusion became possible but not inevitable after the growth of nation-states and their imperial conquests. It required those national, hegemonic and imperial powers to recognize the rights of others, which they were often slow to do. Insofar as morality is treated as universal, it is a result of a diffusion of sympathy, empathy, and morality that came about through growing intercourse and mutual

¹⁶ But their extent is a matter of dispute among psychologists and linguists (Pinker 1994; Deacon 1997; Sampson 2005; Evans and Levinson 2009).

understanding, building on emotions and moral norms that had evolved previously through group selection.

The development of moral dispositions in individuals is necessarily a matter of both inherited value-intuitions and developmental processes within a social culture and community. Morality is more than an individual attribute; hence some moral philosophers refer to *moral systems* (Thiroux and Krasemann 2009; Haidt 2012). These systems include mechanisms of moral inculcation and enforcement in the group. Morality depends on social relations.

This relates to an important point about the nature of socio-economic reality. As Kenneth Arrow (1994) argued forcefully, although economists often make contrary claims, all known economic analyses involve social relations or structures as well as individuals. Hodgson (2007, 2013) shows that “methodological individualism” is a highly ambiguous and lamentably imprecise doctrine, and plausible versions treat social relations as well as individuals as part of the *explanantia*. Socio-economic systems comprise not simply individuals but also relations between them.

Accordingly, morality cannot be understood *simply* as individual attributes or dispositions, however complex or multidimensional. It cannot be encompassed adequately by the “social” or “other-regarding” preferences of an individual, or by a “multiple self” with plural preference functions (Margolis 1982; Elster 1986). While individuals and their attributes must be part of the explanation, we also have to bring in the role of structured relations between individuals. Furthermore, morality is both a biological and a social phenomenon, and is irreducible to either individual preferences or genetic endowments.

Moral systems depend on social positions and relations of authority. Morality is sustained not merely through genetic and cultural transmission, but through the replication of structures of authority and power. According to Haidt and Joseph (2008), respect for authority is one of our five evolved value-intuitions.¹⁷ It has coevolved with these structures, so together they provide powerful systems of social cohesion, as well as sources of inequality and oppression. Hence what is involved is not only imitation or learning, but also the reproduction of institutions and cultural mechanisms for legitimating positions of authority (Runciman 2002).

The evolution and sustenance of moral systems involves a battle between the selfish inclinations of individuals and group moral pressure to conform and cooperate. If the group pressure is insufficient, then individual fitness will be enhanced by acting selfishly. If the group moral pressure restrains selfishness sufficiently, then individual fitness will be enhanced by cooperating with others. It is a classic case of individual fitness being context dependent. In such circumstances we can also speak of group fitness.

In sum, morality depends on language, communication of abstract concepts, discussion and reflection upon them, and the derivation of principles that are general (at least for the group). We inherit some core moral capacities genetically, but not morality as a

¹⁷ On the basis of experimental evidence, Haidt (2012) extends this list to six “moral foundations,” namely care/harm, liberty/oppression, fairness/cheating, loyalty/betrayal, authority/subversion, and sanctity/degradation.

whole. The development of morality proper depends on interaction with others in a social context.

6.6 CONCLUSION: ECONOMICS WITHOUT ECONOMIC MAN

Morality, as understood by leading moral philosophers, cannot be incorporated into models based on individual utility maximization, even with “social” or “other-regarding” preferences. This is because morality is about “doing the right thing,” even if it would otherwise not be the preferred option. Moral judgments are inescapable and cannot be reduced to mere preferences. I argue elsewhere that any form of behavior can be fitted into a utility function. But fitting behavior into functions does not explain its cause or motivation.¹⁸ Digging deeper into the evolutionary and cultural origins of our motives leads us to the issue of morality. Moral motives have evolutionary origins and are sustained through interaction with others: morality is a social as well as an individual phenomenon.

Humans have moral capacities. But we are also self-interested. Evolution has provided us with instincts that trigger our moral development in suitable socio-cultural settings, and with basic instincts such as hunger and lust that can be spurs to egoism. Through our socialization we typically develop into complex personalities where all biologically inherited impulses are extended or constrained to different degrees and in different ways.

The diverse inner impulses that we bring into the world may come into conflict as our personality develops, in the institutional settings of parental care, peer group interaction, and organized education. These settings have major effects on how the moral and self-interested aspects of our personalities develop. Given our declining potential for adaptation as we get older, the earliest years are the most formative.

While accepting that individuals have multi-faceted personalities, influential neo-classical economists assumed that in the economic sphere self-interest was overwhelming, and our altruistic and moral tendencies could be ignored as we entered the world of contract and business (Jevons 1871; Edgeworth 1881; Wicksteed 1910). Writers such as Gary Becker (1976) claimed that utility maximization, developed in the neoclassical analysis of business life, applies generally to all social interactions.

This view has now been widely challenged, on the basis of experimental or other evidence (Bowles and Gintis 2011; Camerer and Fehr 2006; Fehr and Fischbacher 2002; Fehr and Camerer 2007). Consequently, leading theorists now propose that individual preferences are often altruistic or “other-regarding” even when dealing with economic

¹⁸ See Hodgson (2013, esp. ch.3) and my reply to Gintis and Hebling (Hodgson 2015b). Even apparently inconsistent or intransitive preference rankings can be forced into a utility function, by taking into account that different choices always take place in an at least slightly different context or with different information.

matters. A further and more radical step is to challenge the very idea of a preference function: rather than utility maximization, morality involves affirmation of identity and “doing the right thing” (Sen 1977, 1987; Hodgson 2013).

Even firms and markets are unavoidably infused with moral considerations (Hobson 1929; McCloskey 2006; Zak 2008). These may be countered or developed by example or circumstance. If policy-makers ignore our moral dispositions and concentrate on self-interest alone, then they will threaten the very fabric of a modern market economy (Schumpeter 1942).

The acknowledgement of moral motivation opens a large agenda for economists. It is highly relevant for the theory of the firm (Minkler 2008; Lopes et al. 2009; Hodgson 2013). It has been argued here that morality cannot be reduced to individual preferences or altruism. Economic policy is not just about maximizing satisfaction while ensuring that no one's utility is decreased; it should be about guiding and enhancing our moral dispositions. Especially from an evolutionary perspective, and even in the competitive world of modern business, there is no excuse for ignoring the evolution of moral systems and the moral motivations of economic agents.

REFERENCES

- Alexander, Richard D. 1987. *The Biology of Moral Systems*. New York: Aldine de Gruyter.
- Andreoni, James. 1995. “Cooperation in Public Goods Experiments: Kindness or Confusion?” *American Economic Review* 85: 891–904.
- Arrow, Kenneth J. 1994. “Methodological Individualism and Social Knowledge.” *American Economic Review Papers and Proceedings* 84: 1–9.
- Becker, Gary S. 1976. *The Economic Approach to Human Behavior*. Chicago, IL: University of Chicago Press.
- Ben-Ner, Avner, and Louis Putterman. 2000. “Some Implications of Evolutionary Psychology for the Study of Preferences and Institutions.” *Journal of Economic Behavior and Organization* 43: 91–9.
- Binmore, Kenneth. 1994. *Playing Fair: Game Theory and the Social Contract. Volume 1*. Cambridge, MA: MIT Press.
- Binmore, Kenneth. 1998. *Just Playing: Game Theory and the Social Contract. Volume 2*. Cambridge, MA: MIT Press.
- Bok, Sissela. 1978. *Lying: Moral Choice in Public and Private Life*. New York: Pantheon.
- Bowles, Samuel, Jung-Kyoo Choi, and Astrid Hopfensitz. 2003. “The Co-Evolution of Individual Behaviors and Social Institutions.” *Journal of Theoretical Biology* 223: 135–47.
- Bowles, Samuel, and Herbert Gintis. 2011. *A Cooperative Species: Human Reciprocity and its Evolution*. Princeton, NJ: Princeton University Press.
- Boyd, Robert, Herbert Gintis, Samuel Bowles, and Peter J. Richerson. 2003. “Evolution of Altruistic Punishment.” *Proceedings of the National Academy of Sciences* 100: 3531–5.
- Boyd, Robert, and Peter J. Richerson. 1985. *Culture and the Evolutionary Process*. Chicago, IL: University of Chicago Press.
- Boyd, Robert, and Peter J. Richerson. 1992. “Punishment Allows the Evolution of Cooperation (or Anything Else) in Sizable Groups.” *Ethology and Sociobiology* 13: 171–95.
- Brown, Donald E. 1991. *Human Universals*. New York: McGraw-Hill.

- Buchanan, Allen, and Russell Powell. 2015. "The Limits of Evolutionary Explanations of Morality and their Implications for Moral Progress." *Ethics* 126: 37–67.
- Camerer, Colin F., and Ernst Fehr. 2006. "When Does 'Economic Man' Dominate Social Behavior?" *Science* 311: 47–52.
- Carpenter, Jeffrey P., and Peter Hans Matthews. 2009. "What Norms Trigger Punishment?" *Experimental Economics* 12: 272–88.
- Carpenter, Jeffrey P., Peter Hans Matthews, and Okomboli Ong'ong'a. 2004. "Why Punish? Social Reciprocity and the Enforcement of Prosocial Norms." *Journal of Evolutionary Economics* 14: 407–29.
- Darwin, Charles R. 1871. *The Descent of Man, and Selection in Relation to Sex*, 2 vols. London: Murray, and New York: Hill.
- Dawkins, Richard. 1976. *The Selfish Gene*. Oxford: Oxford University Press.
- De Waal, Frans B. M. 1996. *Good Natured: The Origin of Right and Wrong in Humans and Other Animals*. Cambridge, MA: Harvard University Press.
- De Waal, Frans B. M. 2006. *Primates and Philosophers: How Morality Evolved*. Princeton, NJ: Princeton University Press.
- Deacon, Terrence W. 1997. *The Symbolic Species: The Co-Evolution of Language and the Brain*. New York: Norton.
- Diamond, Jared. 1991. *The Rise and Fall of the Third Chimpanzee*. London: Vintage.
- Durham, William H. 1991. *Coevolution: Genes, Culture, and Human Diversity*. Stanford, CA: Stanford University Press.
- Edgeworth, Francis Y. 1881. *Mathematical Psychics: An Essay on the Application of Mathematics to the Moral Sciences*. London: Kegan Paul.
- Etzioni, Amitai. 1988. *The Moral Dimension: Toward a New Economics*. New York: Free Press.
- Evensky, Jerry. 2005. "Adam Smith's Theory of Moral Sentiments: On Morals and Why They Matter to a Liberal Society of Free People and Free Markets." *Journal of Economic Perspectives* 19: 109–30.
- Elster, Jon. 1986. *The Multiple Self*. Cambridge: Cambridge University Press.
- Evans, Nicholas, and Stephen C. Levinson. 2009. "The Myth of Language Universals: Language Diversity and its Importance for Cognitive Science." *Behavioral and Brain Sciences* 32: 429–92.
- Fehr, Ernst, and Colin F. Camerer. 2007. "Social Neuroeconomics: The Neural Circuitry of Social Preferences." *Trends in Cognitive Science* 11: 419–27.
- Fehr, Ernst, and Urs Fischbacher. 2002. "Why Do Social Preferences Matter—The Impact of Non-Selfish Motives on Competition, Cooperation and Incentives." *Economic Journal* 112: C1–C33.
- Fehr, Ernst, and Simon Gächter. 2000a. "Fairness and Retaliation: The Economics of Reciprocity." *Journal of Economic Perspectives* 14: 159–81.
- Fehr, Ernst, and Simon Gächter. 2000b. "Cooperation and Punishment in Public Goods Experiments." *American Economic Review* 90: 980–95.
- Fehr, Ernst, and Simon Gächter. 2002. "Altruistic Punishment in Humans." *Nature* 415: 137–40.
- Fehr, Ernst, and Herbert Gintis. 2007. "Human Motivation and Social Cooperation: Experimental and Analytical Foundations." *Annual Review of Sociology* 33: 43–64.
- Field, Alexander J. 2001. *Altruistically Inclined? The Behavioral Sciences, Evolutionary Theory, and the Origins of Reciprocity*. Ann Arbor, MI: University of Michigan Press.
- Field, Alexander J. 2007. "Beyond Foraging: Evolutionary Theory, Institutional Variation, and Economic Performance." *Journal of Institutional Economics* 3: 265–91.

- Fiske, Alan Page. 1991. *Structures of Social Life: The Four Elementary Forms of Human Relations: Communal Sharing, Authority Ranking, Equality Matching, Market Pricing*. New York: Free Press.
- Frey, Bruno S. 1997. *Not Just for Money: An Economic Theory of Personal Motivation*. Aldershot, UK: Edward Elgar.
- Ghiselin, Michael T. 1974. *The Economy of Nature and the Evolution of Sex*. Berkeley, CA: University of California Press.
- Gintis, Herbert. 2000. "Strong Reciprocity and Human Sociality." *Journal of Theoretical Biology* 206: 169–79.
- Gintis, Herbert, Samuel Bowles, Robert Boyd, and Ernst Fehr (eds.). 2005. *Moral Sentiments and Material Interests: The Foundations of Cooperation in Economic Life*. Cambridge, MA: MIT Press.
- Gintis, Herbert, and Dirk Helbing. 2015. "Homo Socialis: An Analytical Core for Sociological Theory." *Review of Behavioral Economics* 2: 1–59.
- Greene, J. D., and Jonathan Haidt. 2002. "How (and Where) Does Moral Judgement Work?" *Trends in Cognitive Sciences* 6: 517–23.
- Guzmán, Ricardo Andrés, Carlos Rodríguez-Sicken, and Robert Rowthorn. 2007. "When in Rome, do as the Romans do: The Coevolution of Altruistic Punishment, Conformist Learning, and Cooperation." *Evolution and Human Behavior* 28: 112–17.
- Güth, Werner. 1995. "On Ultimatum Bargaining Experiments – A Personal Review." *Journal of Economic Behavior* 27: 329–44.
- Haidt, Jonathan. 2001. "The Emotional Dog and the Rational Tail: A Social Interactionist Approach to Moral Judgment." *Psychological Review* 108: 814–34.
- Haidt, Jonathan. 2012. *The Righteous Mind: Why Good People are Divided by Politics and Religion*. London: Penguin.
- Haidt, Jonathan, and Craig Joseph. 2004. "Intuitive Ethics: How Innately Prepared Intuitions Generate Culturally Variable Virtues." *Daedalus* 133 (4): 55–66.
- Haidt, Jonathan and Joseph, Craig (2008) "The moral mind: How five sets of innate intuitions guide the development of many culture-specific virtues, and perhaps even modules." in Peter Carruthers, S. Laurence, and S. Stich (eds.) *The Innate Mind, Vol. 3*, pp. 367–444.
- Hamilton, William D. 1964. "The Genetical Evolution of Social Behavior." *Journal of Theoretical Biology* 7: 1–32.
- Hamlin, J. Kiley, Karen Wynn, and Paul Bloom. 2007. "Social Evaluation by Preverbal Infants." *Nature* 450: 557–60.
- Hammerstein, Peter (ed.). 2003. *Genetic and Cultural Evolution of Cooperation*. Cambridge, MA: MIT Press.
- Hare, Richard M. 1952. *The Language of Morals*. Oxford: Oxford University Press.
- Hauser, Marc D. 2006. *Moral Minds: How Nature Designed Our Universal Sense of Right and Wrong*. New York: Harper Collins.
- Henrich, Joseph. 2004. "Cultural Group Selection, Coevolutionary Processes and Large-Scale Cooperation." *Journal of Economic Behavior and Organization* 53: 3–35.
- Henrich, Joseph, and Robert Boyd. 1998. "The Evolution of Conformist Transmission and the Emergence of Between Group Differences." *Evolution and Human Behavior* 19: 215–42.
- Henrich, Joseph, Robert Boyd, Samuel Bowles, Colin Camerer, Ernst Fehr, and Herbert Gintis. 2004. *Foundations of Human Sociality: Economic Experiments and Ethnographic Evidence from Fifteen Small-Scale Societies*. Oxford and New York: Oxford University Press.

- Henrich, Joseph, Robert Boyd, Samuel Bowles, Colin Camerer, Ernst Fehr, Herbert Gintis, and Richard McElreath. 2001. "In Search of Homo Economicus: Behavioral Experiments in 15 Small-Scale Societies." *American Economic Review Papers and Proceedings* 91: 73–84.
- Henrich, Joseph, Jean Ensminger, Richard McElreath, et al. 2010. "Markets, Religion, Community Size, and the Evolution of Fairness and Punishment." *Science* 327: 1480–4.
- Henrich, Joseph, and Francisco J. Gil-White. (2001) "The Evolution of Prestige: Freely Conferred Deference as a Mechanism for Enhancing the Benefits of Cultural Transmission." *Evolution and Human Behavior* 22: 165–96.
- Henrich, Joseph, Richard McElreath, Abigail Barr, et al. 2006. "Costly Punishment Across Human Societies." *Science* 312: 1767–70.
- Hill, Kim R., Robert S. Walker, Miran Božičević, James Eder, Thomas Headland, Barry Hewlett, et al. 2011. "Co-Residence Patterns in Hunter-Gatherer Societies Show Unique Human Social Structure." *Science* 331: 1286–9.
- Hobson, John A. 1929. *Wealth and Life: A Study in Values*. London: Macmillan.
- Hodgson, Geoffrey M. 1993. *Economics and Evolution: Bringing Life Back Into Economics*. Cambridge, UK and Ann Arbor, MI: Polity Press and University of Michigan Press.
- Hodgson, Geoffrey M. 2007. "Meanings of Methodological Individualism." *Journal of Economic Methodology* 14: 211–26.
- Hodgson, Geoffrey M. 2009. "On the Institutional Foundations of Law: The Insufficiency of Custom and Private Ordering." *Journal of Economic Issues* 43: 143–66.
- Hodgson, Geoffrey M. 2013. *From Pleasure Machines to Moral Communities: An Evolutionary Economics without Homo Economicus*. Chicago, IL: University of Chicago Press.
- Hodgson, Geoffrey M. 2014. "The Evolution of Morality and the End of Economic Man." *Journal of Evolutionary Economics* 24: 83–106.
- Hodgson, Geoffrey M. 2015a. *Conceptualizing Capitalism: Institutions, Evolution, Future*. Chicago, IL: University of Chicago Press.
- Hodgson, Geoffrey M. 2015b. "A Trojan Horse for Sociology? Preferences versus Evolution and Morality." *Review of Behavioral Economics* 2: 93–112.
- Hodgson, Geoffrey M., and Knudsen, Thorbjørn. 2010. *Darwin's Conjecture: The Search for General Principles of Social and Economic Evolution*. Chicago, IL: University of Chicago Press.
- Huxley, Thomas Henry, and Julian Huxley. 1947. *Evolution and Ethics, 1893–1943*. London: Pilot Press.
- Jevons, William Stanley. 1871. *The Theory of Political Economy*. London: Macmillan.
- Joyce, Richard. 2006. *The Evolution of Morality*. Cambridge, MA: MIT Press.
- Kaplow, Louis, and Steven Shavell. 2007. "Moral Rules, the Moral Sentiments, and Behavior: Toward a Theory of an Optimal Moral System." *Journal of Political Economy* 115: 494–514.
- Kropotkin, Petr A. 1902. *Mutual Aid: A Factor of Evolution*, 1st edn. London: Heinemann.
- Lopes, Helena, Ana C. Santos, and Nuno Teles. 2009. "The Motives for Cooperation in Work Organisations." *Journal of Institutional Economics* 5: 259–88.
- Mackie, John Leslie. 1977. *Ethics: Inventing Right and Wrong*. Harmondsworth: Penguin.
- Margolis, Howard. 1982. *Selfishness, Altruism, and Rationality: A Theory of Social Choice*. Chicago, IL: University of Chicago Press.
- McCloskey, Deirdre N. 2006. *The Bourgeois Virtues: Ethics for an Age of Commerce*. Chicago, IL: Chicago University Press.
- McCrink, Koleen, Paul Bloom, and Laurie R. Santos. 2010. "Children's and Adult's Judgments of Equitable Resource Distributions." *Developmental Science* 13: 37–45.

- Menger, Carl. 1871. *Grundsätze der Volkswirtschaftslehre*. Tübingen: J. C. B. Mohr. Translated and published in English in 1981 as *Principles of Economics* (New York: New York University Press).
- Midgley, Mary. 2003. *Myths We Live By*. London and New York: Routledge.
- Minkler, Lanse P. 2008. *Integrity and Agreement: Economics When Principles Also Matter*. Ann Arbor, MI: University of Michigan Press.
- Nichols, Shaun. 2004. *Sentimental Rules: On the Natural Foundations of Moral Judgment*. Oxford and New York: Oxford University Press.
- Oppenheimer, Stephen. 2004. *The Real Eve*. New York: Carroll and Graf. Published in the UK as *Out of Eden* (London: Constable and Robinson).
- Palmer, Craig T., and Lyle B. Steadman. 1997. "Human Kinship as a Descendant-Leaving Strategy: A Solution to an Evolutionary Puzzle." *Journal of Social and Evolutionary Systems* 20: 39–51.
- Pinker, Steven. 1994. *The Language Instinct: The New Science of Language and Mind*. London and New York: Allen Lane and Morrow.
- Price, Michael E., Leda Cosmides, and John Tooby. 2002. "Punitive Sentiment as an Anti-Free-Rider Device." *Evolution and Human Behavior* 23: 203–31.
- Rizzolatti, Giacomo, Corrado Sinigaglia, and Frances Anderson. 2007. *Mirrors in the Brain: How Our Minds Share Actions, Emotions, and Experience*. Oxford and New York: Oxford University Press.
- Robinson, Paul H., Robert Kurzban, and Owen D. Jones. 2007. "The Origins of Shared Intuitions of Justice." *Vanderbilt Law Review* 60: 1633–88.
- Roberts, Simon. 1979. *Order and Discipline: An Introduction to Legal Anthropology*. New York: St Martin's Press.
- Runciman, Walter Garry. 2002. "Heritable Variation and Competitive Selection as the Mechanism of Sociocultural Evolution." In, Michael Wheeler, John Ziman, and Margaret A. Boden (eds.), *The Evolution of Cultural Entities* (Oxford: Oxford University Press), pp. 9–25.
- Sampson, Geoffrey. 2005. *The Language Instinct Debate*. London and New York: Continuum.
- Schumpeter, Joseph A. 1942. *Capitalism, Socialism and Democracy*, 1st edn. London: George Allen and Unwin.
- Schwartz, Shalom H. 1994. "Are There Universal Aspects in the Structure and Contents of Human Values?" *Journal of Social Issues* 50: 19–45.
- Sen, Amartya K. 1977. "Rational Fools: A Critique of the Behavioral Foundations of Economic Theory." *Philosophy & Public Affairs* 6: 317–44.
- Sen, Amartya K. 1987. *On Ethics and Economics*. Oxford and New York: Basil Blackwell.
- Smart, J. J. C., and Bernard Williams. 1973. *Utilitarianism: For and Against*. Cambridge: Cambridge University Press.
- Smith, Adam. 1759. *The Theory of Moral Sentiments*. London and Edinburgh: Millar, and Kincaid and Bell.
- Smith, Adam. 1776. *An Inquiry into the Nature and Causes of the Wealth of Nations*, 2 vols. London: Strahan and Cadell.
- Sober, Elliott, and David Sloan Wilson. 1998. *Unto Others: The Evolution and Psychology of Unselfish Behavior*. Cambridge, MA: Harvard University Press.
- Tancredi, Laurence. 2005. *Hardwired Behavior: What Neuroscience Reveals about Morality*. Cambridge and New York: Cambridge University Press.

- Thiroux, Jacques P., and Keith W. Krasemann. 2009. *Ethics: Theory and Practice*, 10th edn. London: Pearson.
- Tisak, Marie S., and Elliot Turiel. 1984. "Children's Conceptions of Moral and Prudential Rules." *Child Development* 55: 1030–9.
- Trivers, Robert L. 1971. "The Evolution of Reciprocal Altruism." *Quarterly Review of Biology* 46: 35–57.
- Tyler, Tom R. 1990. *Why People Obey the Law*. New Haven, CT: Yale University Press.
- Vanberg, Viktor J. 2008. "On the Economics of Moral Preferences." *American Journal of Economics and Sociology* 67: 605–28.
- Wade, Michael J. 1978. "A Critical Review of the Models of Group Selection." *Quarterly Review of Biology* 53: 101–14.
- Walras, Léon. 1874. *Éléments d'économie politique pure, ou théorie de la richesse sociale*. Lausanne: Rouge.
- Wiessner, Polly. 2005. "Norm Enforcement among the Ju/'hoansi Bushmen: A Case of Strong Reciprocity?" *Human Nature* 16: 115–45.
- Westermarck, Edward. 1891. *The History of Human Marriage*. London and New York: Macmillan.
- Wicksteed, Philip H. 1910. *The Commonsense of Political Economy, Including a Study of the Human Basis of Economic Law*. London: Macmillan.
- Williams, George C. 1966. *Adaptation and Natural Selection*. Princeton, NJ: Princeton University Press.
- Wilson, David Sloan. 1980. *The Natural Selection of Populations and Communities*. Menlo Park, CA: Benjamin/Cummings.
- Wilson, David Sloan. 1983. "The Group Selection Controversy: History and Current Status." *Annual Review of Ecology and Systematics* 14: 159–88.
- Wilson, David Sloan. 2002. *Darwin's Cathedral: Evolution, Religion, and the Nature of Society*. Chicago, IL: University of Chicago Press.
- Wilson, David Sloan, and Elliot Sober. 1994. "Reintroducing Group Selection to the Human Behavioral Sciences." *Behavioral and Brain Sciences* 17: 585–608.
- Wilson, David Sloan, and Edward O. Wilson. 2007. "Rethinking the Theoretical Foundations of Sociobiology." *Quarterly Review of Biology* 82: 327–48.
- Zak, Paul J. 2004. "Neuroeconomics." *Philosophical Transactions B of the Royal Society of London* 359: 1737–48.
- Zak, Paul J. (ed.) 2008. *Moral Markets: The Critical Role of Values in the Economy*. Princeton, NJ: Princeton University Press.

CHAPTER 7

MORALITY AS A COMPLEX ADAPTIVE SYSTEM

Rethinking Hayek's Social Ethics

GERALD GAUS

7.1 PHILOSOPHY AND ECONOMICS: INTERDISCIPLINARY AND TRANSDISCIPLINARY ANALYSES

DANIEL M. HAUSMAN and Michael S. McPherson commence their highly influential *Economic Analysis, Moral Philosophy and Public Policy* by providing a classic specification of the scope of the study of ethics and economics:

We prescribe no code of conduct and preach few sermons. Rather in this book we try to show how understanding moral philosophy can help economists to do better economics and how economics and ethics can help policy analysts to improve their evaluations of alternative policies. We also hope to show how philosophers can do ethics better by drawing on insights and analytical tools from economists. (2006: 3)

We might call this the cross-fertilization program for *interdisciplinary* inquiry. Each discipline brings its core findings and tools to bear in ways that enlighten and refine the other discipline. The study of “philosophy, ethics, and economics,” then, takes place at the intersection of two distinct, well-developed, disciplines, and learns from each. This interdisciplinary exchange between philosophers and economists has proved immensely fruitful. Economists have, for example, taught welfare-oriented philosophers a great deal about aggregating utility functions (Sen 2017) and political philosophers

about game theory and the dynamics of cooperation (Sugden 1986; Binmore 1994, 1998). And philosophers from Hume (see Hardin 2007) to Rawls (1999) have contributed greatly to the level of sophistication that economists bring to the study of welfare, fairness, justice, and rights.

In recent years some have embarked on *transdisciplinary* analyses (Page 2018). While this term is employed in many ways (some perhaps regrettable), one core use is to signal that the same model that enlightens us in one field applies equally well to a very different discipline's area of study. In contrast to interdisciplinarity, we are not sharing the results of two distinct disciplines, but seeing that what at first looked like very different phenomena have the same underlying dynamics. A complexity model from physics, for instance, might straightforwardly apply in biology, or an analysis of biological selection may apply to organizational theory or scientific inquiry (Weisburg and Muldoon 2009; D'Agostino 2010). Or, as I shall try to show in this chapter, models of complex adaptive systems apply to moral norms as well as to the market. Both, I shall argue, are self-sustaining spontaneous orders (Robbins 1976: 9). Or, rather, both are part of the same adaptive order.

Though “transdisciplinarity” is something of a buzz word today, the idea is not new. Throughout his long career, and especially from the 1950s through the 1970s, F.A. Hayek repeatedly stressed that the model of a complex evolutionary order characterized social evolution, the economic order, and morality itself. For half a century Hayek was essentially a lone voice in the wilderness. While economists became increasingly obsessed with highly abstract, usually static, mathematical models, and philosophers got lost in conceptual and utopian analyses—with almost everyone abhorring the very idea of social evolution¹—Hayek stressed that economic systems were constantly evolving, adapting, and dependent on an unplanned moral framework, which was also constantly adapting (Gaus 2013b). Only recently have some economists, philosophers, and complex systems scientists recaptured this insight. In recent work, biologists, anthropologists, political scientists, philosophers, and economists have come to recognize how both biological and social systems form complex adaptive systems, and that social rules, including moral rules, are part of these systems.

This chapter focuses on Hayek's analysis of morality as an evolved spontaneous order, while updating and revising it, taking account of current research and models. We shall see that while, of course, his path-breaking work requires revision, the heart of his analysis is reinforced by recent work in evolutionary theory, complex systems analysis, and social norms theory. Hayek presents an analysis of a complex moral order that is far more in tune with current science than are the highly rationalistic analyses of contemporary political philosophy, which often seek to present utopian plans for the perfectly just society (Gaus 2016). Yet, I shall argue, we need to rethink important claims, as is befitting an ongoing research program.

¹ One of the most pejorative terms of all—“Social Darwinist”—was (quite wrongly) applied to Hayek, by both libertarians (Paul 1988) and by social democrats (Miller 1989). Until the 1990s, the very idea of social evolution was often thought to have fascist overtones.

7.2 RULES

7.2.1 Mind and Rules

Complex systems are rule-governed (Lane 2017: 10). Whereas *Homo economicus* has a difficult time explaining rule-following (Gaus 2011: ch. 2), Hayek puts it at the core of his analysis: “Man is as much a rule-following animal as a purpose-seeking one” (1973: 11). This claim is deeply embedded in his neural network account of the mind (Gaus 2006; Lewis 2015): the mind itself is inherently rule-governed. On connectionist theories such as Hayek’s, the brain is composed purely of neurons that simply have on/off states: qualitative differences (for example, between thoughts) are the results of a complex pattern of neural activation (Hayek 1952: 57). On this account, neurons are either excited or inhibited by other units; the output of the entire network is a function of its initial conditions and the pattern of activation between the neurons. The system changes the weights of the excitatory and inhibitory inputs according to some learning rule, so inputs are represented simply by the changes in weights they have caused (Gärdenfors 2004: 41).

The subject of Hayek’s neural network theory is, as he explains, “the kind of process by which a given physical situation is transformed into a certain phenomenal picture” (1952: 7). A certain state of the external world W exists at time t : how is W_t transformed into a sensory experience S of W_t , and how does $S[W_t]$ relate to sensory experiences of other states of the world—when will these be perceived as the same as $S[W_t]$, and when will the sensation be different? The key to Hayek’s analysis is “classification” via neuronal connections, “a process of channeling, or switching, or ‘gating,’ of the nervous impulses so as to produce a particular disposition or set” (1967b: 51). As he elaborates:

By “classification” we shall mean a process in which on each occasion on which a certain recurring event happens it produces the same specific effect, and where the effects produced by any one kind of such events may be either the same or different from those which any other kind of event produces in a similar manner. All the different events which whenever they occur produce the same effect will be said to be events in the same class, and the fact that every one of them produces the same effect will be the *sole* criterion which makes them members of the same class. (1952: 48)

Thus two events are the same just in case they trigger the same neuronal configuration. The central nervous system, then, takes what we might think of as an undifferentiated world and, via the connections in the neuronal network, creates a formal structure of classes of sensations (Hayek 1952: 51). But this is to say that the mind is rule-governed: the neuronal connections constitute perceptions of patterns. In a fundamental sense the mind *is* a set of rules that takes sensory inputs and yields classifications and perceptions.

Cristina Bicchieri’s (2016) influential account of social norms is also grounded in a connectionist theory of the mind (Bicchieri and McNally 2018). On her analysis a

norm or social rule can be understood as involving a sort of scripted activity. When an individual categorizes a situation as, say, a commercial exchange, a sort of script is invoked about payment and relevant information. But before the script can be invoked individuals must have categorized the situation. Schemata, on Bicchieri and McNally's account,

are knowledge structures that help people interpret the world around them. The more elements of a schema that we observe, the more likely that schema will be activated. The fewer elements relevant to a schema that one observes or the less prototypical the elements are, the less likely the schema will be activated. A prototype is the 'standard' conceptualization of a particular kind, category, or phenomenon. (2018)

For example:

Depending on how one interprets the situation, subjects could either decide that reciprocation or equality is more important, ultimately resulting in completely different decisions. In this respect, changing these schematic lenses can serve to alter how norms are activated. . . .

As is implied by the connectionist underpinnings of schema theory, individual scripts and schemata do not exist in isolation; they are inherently linked to each other to varying degrees, and the activation of one influences the activation of another. The entirety of one's interconnected schemata is termed a semantic or associative network. A semantic network is a model of conceptual interconnectivity, with each individual schema serving as a node, and each relationship between schemata represented as a link of varying strengths. Chronic activation of multiple schemata in tandem will increase the strength of their associative links. (Ibid.)

Schemata, which provide a general system of categories that allow us to interpret and classify social situations, and scripts, which instruct us what a social rule calls for in these situations, are produced by reinforcement of connections within the semantic network.

7.2.2 Rules and Action

In a fundamental sense, then, rules are prior to purposes. In order to even characterize a situation as one relevant to goal pursuit—for example, “this is a case of profit-to-be-had”—the agent must already have applied a system of categorizations (Gaus 2011: 144–5). An agent's schemata must have identified this as counting as a case where profit-making is relevant. Given this, the popular project of deriving rule-following from goal pursuit—one that has been especially attractive to theorists impressed by economic models (Gaus 2011: ch. 2, 2018a)—gets things almost exactly backwards, at least as far as the

workings of the mind are concerned.² And for Hayek, the fundamental place of rule following in cognition carries over into action or “motor behavior.” Action also follows from neural connections and so human action is inherently rule-based: it is regulated by a network that is based on classification of types of sensation and how they relate to types of responses. Practical rules identify “patterns of actions” that are classified as having the same meaning; the activation of such rules disposes the agent to act (Hayek 1967b: 57).

Hayek is surely right that the social life of humans, from its earliest moments, developed within a framework of action-guiding social rules.³ Indeed, human genetic evolution has for long periods occurred within culture and its moral rules, such that the environment in which genetic evolution has taken place has been rule-governed.⁴ In this sense, it is entirely appropriate to say that we are, by nature, rule-following creatures (Gaus 2011: ch. 3). As Henrich stresses, “While many evolution and economics oriented researchers have often assumed that social norms...are merely a superficial window dressing on our evolved psychology, the evidence suggests that such social norms run deep and profoundly shape social life” (2016: 149). It is thus a fundamental error of the common economic mindset to suppose that a “mere norm is unlikely to override self-interest in many...contexts” (Hardin 2013: 414). Experimental and field evidence has consistently shown the power of norms or social rules to shape behavior (Bicchieri 2016; Gaus 2018b).

7.3 COMPLEXITY

7.3.1 Bodo Rules

For much of our evolutionary history the evidence indicates that humans lived in relatively small groups of anywhere between 25 and 150 persons.⁵ Indeed, for most of recorded history most humans lived in small face-to-face groups with low levels of social diversity. There were, of course, differences in gender roles, age, sometimes modest occupational differences and, of course, expected personality variance.⁶ In these simple societies, the rules of social life could cover most eventualities: the useful or appropriate act-types could be identified by the group’s history. Individual invention

² “In humans you first need to acquire the social norms and rules governing the world you are operating in, and only then is strategic thinking useful” (Henrich 2016: 52).

³ See Boehm (2012: 217ff), Kitcher (2011: ch. 2), and Gaus (2011: ch. 3, 2018b). In this essay I do not distinguish social rules, social norms, and the rules of social morality. While in some contexts these are fundamentally different, in this case nothing significant turns on their differences.

⁴ See Boehm (2012), Richerson and Boyd (2005), and Henrich (2016).

⁵ Friedman (2008: 16) points to 150, with much larger numbers when groups fused. See also Rose (2011: ch. 3), who mentions 200 as the typical size of the groups in which humans evolved. Closer examination shows that group size may be understood differently: average band size may differ from typical group size (Bowles and Gintis 2011: 95).

⁶ Ethnographic accounts of small-scale societies wonderfully reveal these. Turnbull’s (1963: ch. 5) tale of the selfish and domineering Cephu is classic.

of “new ways of living” were at best infrequent and gradual (Hardin 2013; Gaus and Nichols 2017).⁷ Moreover, because of the homogeneity in metaphysical, religious, moral, and cultural values (see, for example, Brandt 1954), how one person acted within a rule would be largely the way his fellows did, so the overall results would be highly predictable.

This is nicely brought out in Russell Hardin’s tale of Bodo:

Axel Leijonhufvud . . . characterizes the village society of eleventh century France in which the villager Bodo lived. We have detailed knowledge of that society from the parish records of the church of St. Germaine. Today one would say that that church is in the center of Paris, but in Bodo’s time it was a rural parish distant enough from Paris that many of its inhabitants may never have seen Paris. Virtually everything Bodo consumed was produced by about eighty people, all of whom he knew well. Indeed, most of what he consumed was most likely produced by his own family. If anyone other than these eighty people touched anything he consumed, it was salt, which would have come from the ocean and would have passed through many hands on the way to St. Germaine, or it was spices, which would have traveled enormous distances and passed through even more hands. (1999: 401–2)

In this world, Hardin observes, the ethics of Bodo’s community probably had some notion of “distributive justice” in the form of

a principle of establishing a lower bound on how impoverished a person could be. . . . In Bodo’s society, charity might be subject to a principle of fairness—not fairness toward the poor but fairness toward one’s peers in shouldering a share of the burden of charity. Such fairness can be assessed in such a community because so much about everyone is a matter of common knowledge. (Ibid.: 403)

Not only did everyone know a great deal about others, but it would be highly predictable how much charity would be given and by whom, how the impoverished would react to charity, as well as the attitudes of those who must bear the burden. The shared cultural outlook not only recognizes the rule, but provides highly regularized ways of complying and responses to compliance.⁸

7.3.2 Large-Scale Society: The Diversity of Compliance Behavior

Because Bodo’s society would have only a modest number of rules, with predictable ways of complying, the rule system would generally have highly predictable consequences. Fundamental to Hayek’s analysis was that social morality undergoes a fundamental

⁷ In the extraordinarily harsh late-Pleistocene climate, change often came much more swiftly (Richerson and Boyd 2013).

⁸ We should not press this point too hard. Research on the Aché distinguishes four types of individuals in terms of their attitudes towards helping others and group production (Boehm 2012: 294–7). The important point is the severe limits of diversity, not the absence of difference.

transformation in the move to large-scale social orders; indeed, he suggests reserving the term “morality” for the “new and different morality” that arises with large-scale social organization (1988: 12). The effects of large numbers and greatly increased heterogeneity of agents combine to produce a qualitative transformation in the nature of a rule-based moral order. Two changes drive this transformation: (i) the increased heterogeneity of the actors complying with the rules and (ii) the increase in the number of rules.

Heterogeneity. Consider the charity rule suggested by Hardin: the rule says that one ought to assist those in severe need. As I remarked, in Bodo’s society, there would be a high degree of consensus on what was needed and who qualified, as well as attitudes about the appropriate responses to giving and receiving charity (allowing for general human variance between more and less altruistic individuals). Now assume that we have a society that still endorses the rule, but agents have a wide variety of different moral perspectives and background theories. For example, some may have a theory that sharply distinguishes deserving and undeserving poor while some may be Malthusians who hold that charity should be linked to reduced opportunities for reproduction. In nineteenth-century England, both of these perspectives might agree that poor relief should take the form of the workhouse rather than outdoor relief, but stress different aspects of the workhouse (compelled labor for the former, segregation of the sexes for the latter) as its morally important feature.⁹ Others rejected the workhouse as inconsistent with true Christian charity or with genuine concern for the poor, who were the victims of industrialization. Given this background diversity, predicting the effects of even universal compliance with a rule of charity becomes exponentially more difficult (the workhouse, outdoor relief, or socialism?), for each different perspective will have different ways of complying, with vastly different effects.

Types of Rules and the Eligible Moral Space. It may be thought that this is simply a problem with vague rules such as “give charity to the very needy,” which allow open-ended interpretations. However, vagueness is not the crux of the matter: almost all rules specify a space or range of compliant action, not a determinative complaint action. Shaun Nichols and I (2017) studied two types of rule systems: permissive and prohibitory. Permissive systems instruct the agent what she is allowed to do. We found that if agents are taught rule systems containing only permission rules, they infer a closure rule for uncovered cases: what is not permitted by one of the rules is prohibited. That is, one is prohibited from acting unless one has a permission stated in some rule. This is in many ways a very restrictive rule system, seeking to guide all actions into permitted channels. Yet even these rule systems allow tremendous diversity of responses. Chess, after all, is a system of permissions: what the rules of chess do not allow is prohibited. One may not say: “Ah! The rules of chess say nothing about melting your queen, so there she goes!” If it is not in the rules, it is disallowed. But the number of possible moves in chess has been estimated as at least 10^{120} .

Nevertheless, pure systems of permissions are in a way especially restrictive: they necessarily identify a set of acceptable act-types, and given that people infer that what is

⁹ Compare Malthus (1798) and Bosanquet (1890).

not permitted is also prohibited, it follows that they are interpreted as disallowing new act-types: act types that are not permitted by the rules are prohibited. Hayek held that in large-scale societies with dispersed information, rules should aim to provide some definite expectations of the actions of others while allowing each individual to pursue her individual plans, given her different values and background assumptions. The critical function of social rules is to secure cooperation while leaving people free to pursue their diverse aims:

One of the main aims of the rules... must therefore be to eliminate for the individual as much avoidable uncertainty as possible. This means that he must be able to ascertain from the circumstances which he can know what he is free to do, and under what circumstances and in what manner other human forces will constrain him. If he is to use his knowledge to the best advantage for achieving his aims, the world around him must be, as far as possible, given to him. Of course, in a changing world much of his task and his merit will be to foresee changes correctly, to adapt himself successfully to ever changing conditions. (Hayek 1955: 162)

Adam Smith held that this result is best secured through the “system of natural liberty” in which “[e]very man, as long as he does not violate the laws of justice, is left perfectly free to pursue his own interest his own way” (1776: 687). For him, justice was largely a matter of negative rules:

Mere justice is, upon most occasions, but a negative virtue, and only hinders us from hurting our neighbour. The man who barely abstains from violating either the person, or the estate, or the reputation of his neighbours, has surely very little positive merit. He fulfils, however, all the rules of what is peculiarly called justice, and does every thing which his equals can with propriety force him to do, or which they can punish him for not doing. We may often fulfil all the rules of justice by sitting still and doing nothing. (1759: 82)

Smith’s system of natural liberty is psychologically coherent and easy to learn. In the study with Nichols, we found that those taught purely prohibition rules, or even sets of rules mixed with prohibitions and permissions, tended to import the principle of natural liberty: if the negative rules do not prohibit an action, it is allowed. Note that systems of natural liberty allow agents to invent new ways of acting: an act-type that is unmentioned by the rules is allowed. Because of this, Smith’s system of natural liberty allows tremendous variation in compliance behavior: there are infinitely many ways *not* to do something (cf. Hayek 1960: 150–2). As the diversity of plans, values and background theories multiplies, the variety of compliance behavior greatly expands in unpredictable ways.

7.3.3 Large-Scale Society: The Size of the Rule System

Because the system of natural liberty allows exploration of new act-types (Gaus and Nichols 2017), as exploration and invention proceed some of these will be found to pose obvious dangers, and so new moral rules will be required to regulate them. Compared to

Bodo's society, in the large and diverse societies encouraged by the system of natural liberty, cultural values and religious and metaphysical theories provide far less reliable social coordination (as they are not deeply shared), again resulting in the growth of the rules of social morality. As the rule system grows, the ways that one rule affects the overall relations among people will be dependent on other rules in the system and the ways that people comply with them (Hayek 1967d: 91). A rule that invokes heavy penalties for stealing may have very different consequences if there is also a rule that requires assistance to those in need, or if the rules of property are vague. This is not a philosophical conjecture. In her extensive fieldwork on actual institutions Elinor Ostrom stressed that institutions are composed of numerous rule configurations; the constituent rules have strong interdependencies, both with each other and with environmental conditions. "A change in any one of these variables produces a different action situation and may lead to very different outcomes" (Ostrom 1986: 111; see also Lewis 2017).

We need not assume maximal interconnectedness of each rule with all others: clusters of rules will form systems reasonably independent of the rest of the set. But identifying these clusters will itself be a matter of discovery. Such interdependencies are seldom formal parts of the rules; a rule's interactions with other rules is often unexpected and can change as circumstances vary. We could not have anticipated, for example, that in response to a law prohibiting the consumption and sale of alcohol, conjoined with certain policing rules and practices, many became much more skeptical of legal regulation and of the police while less suspicious of organized crime (Stuntz 2000). As individuals witness each other's behavior under the system of rules, each takes her observations into account in her own future compliance behavior. Observing, for example, that the prohibition of alcohol was most systematically directed at the poorer, most easily monitored beer-drinking population, an agent may change her consumption to spirits, begin to draw back cooperation with the police, or become less suspicious of her suppliers, organized crime (*ibid.*). As some individuals change their behavior in these ways, others will in turn adjust theirs. At each point, the compliant behavior of some is an input into new compliance decisions of others, perhaps inducing non-compliance. In more technical terms, the system is one with multiple feedback loops. It may well be impossible to predict the future behavior of such orders: systems of equations with multiple levels of feedback quickly become incalculable.

7.3.4 Complexity and Emergence

As diversity and scale increase, the overall effects of the rule system thus become complex and perhaps impossible to predict in any detail. Once we have arrived at this juncture, the overall moral order—how individuals morally relate to each other in a society—is an *emergent property* of countless diverse individuals responding to each other's moral decisions. Having built in widespread diversity, considerable interconnectedness,

and individual moral choices according to rules, the moral system is, in a technical sense, complex.¹⁰ Moreover, if we think of social morality as a technology of cooperation (Kitcher 2011: 221–41; Gaus 2013a), then like other technologies, it seems to become increasingly complex as it develops (Arthur 2015: ch. 1). Rules are added as refinements on the preexisting system, producing yet more interconnections and, so more complexity.

Hayek (1967b, 1967c) was one of the first to recognize that under these conditions of complexity, the overall moral order will be an emergent property of the system of rules employed by diverse agents (Lewis 2015, 2017). Emergent properties are sometimes distinguished from mere “resultant” properties on the grounds that, while a resultant property is the expected consequence of an underlying set of properties, emergent properties are very often novel and unexpected. In perhaps the earliest analysis of such systems, John Stuart Mill (1843: 370–3, 438ff; see also Auyang 1998: 173ff) considered a system, *S*, composed of elements (such as rules) $\{r_1 \dots r_n\}$, and an overall resulting order *O*. Mill proposes three features of property *O*:

- (1) *O* is not the sum of $\{r_1 \dots r_n\}$;
- (2) *O* is of an entirely different character than $\{r_1 \dots r_n\}$;
- (3) *O* cannot be predicted or deduced from the behavior of the members of $\{r_1 \dots r_n\}$ considered independently (that is, apart from their interactions in *S*).

Thus it is said that waves are an emergent property of H_2O . The properties studied by hydrology are not the sum of the properties of hydrogen and oxygen (as opposed to a mechanical force, which may be seen as the sum of its causes); waves have a very different character than a chemical compound, and the properties of an individual water molecule do not allow us to deduce the relevant laws concerning waves.

The scientific study of complex orders cannot, Hayek (1967b) rightly insisted, aim at the prediction of the “specific” future states or values of the individual elements. To be sure, just how specific a specific prediction needs to be is context-dependent; his claim, though, was that in many natural sciences (such as parts of physics), “it will generally be possible to specify all those aspects of the phenomenon in which we are interested with any degree of precision we may need for our purposes” (1967a: 8). In contrast, when dealing with complex phenomena we are simply unable to specify the values. In contemporary terms, the system is modeled in nonlinear equations (Vaughn 1999: 245; Holland 2014: 4; Mitchell 2009: 22ff.): we can only predict the “range of phenomena to expect” (Hayek 1967a: 11, 1967b). We can understand the general principles on which the system operates, and with this knowledge we can predict the parameters within which the system will settle (Hayek 1955: 43).

¹⁰ On complexity, see Mitchell (2009), Page (2011), and Holland (2014).

7.4 COMPLEX SYSTEMS: MACRO ADAPTATION

7.4.1 Macro Selection

Once we examine their features, it is manifest that systems of moral rules in large-scale diverse societies are complex. The concept of a *complex adaptive system* identifies complex systems that arise through agents responding to each other's action, and adapting their actions in response to each other (Holland 2014: ch. 3; Miller and Page 2007). In an important essay, D.S. Wilson (2016) presses the question of what renders these complex interactions *adaptive* as opposed to simply complex. Great complexity can lead to "complexity catastrophe" which results in chaos (Kauffman 1993: 52–4). Wilson distinguishes two mechanisms of adaptation: selection pressures at the system or macro-level (discussed in this section) and rational adaptation of each agent (discussed in the next).

What Wilson refers to under the (less than memorable) designation of "CAS₁" systems, the complex system is subject to adaptive pressures at the system-level. This is a form of multi-level selection (Okasha 2006): in the most familiar form, a type of "group selection." On Wilson's view this is the most plausible basis for system-level adaptation: "[f]rom an evolutionary perspective . . . only when a society is a unit of selection . . . does it function well as a unit" (2016: 44). The critical claim is that a complex social order will maintain its cooperative functionality only if, at the societal level, forces are constantly selecting more over less functional variants. And this, perhaps surprisingly, was precisely Hayek's view. At the macro level, what is selected, Hayek argues (1967c: 71), is an "order of actions," the emergent property of moral order that arises in a rule-based society. At the macro level evolutionary selection pressures operate directly on "the order of actions of a group" (ibid.: 72). This distinction between a set of rules and the order of actions to which it gives rise is a fundamental insight of Hayek's, which allows him to distinguish the focus of selective pressure (the functionality of the order) and the underlying rules (that structure it), which are transmitted. On Hayek's analysis (1967b: 23–4), a group of individuals living under a set of social rules R , composed of rules $\{r_1 \dots r_n\}$, will give rise to a certain emergent pattern of social interactions, O , on which macro selection operates: "systems of rules of conduct will develop as a whole" (Hayek 1967d: 71).

On Hayek's analysis, macro social evolution is a form of cultural group selection: "The rules of conduct have . . . evolved because the groups who practiced them were more successful and displaced others" (1973: 18). If society S_1 , characterized by order of actions O_1 , is more productive than S_2 based on O_2 , society S_1 will tend to win conflicts with S_2 , a mechanism akin to natural selection: O_1 was more adaptive than O_2 (Bowles and Gintis 2011). But perhaps more importantly, the members of S_2 , seeing the better-off participants in S_1 characterized by O_1 , may either immigrate to S_1 , or seek to copy the underlying rules R_1 , thus inducing differential rates of reproduction between the two sets of underlying rules (Richerson and Boyd 2005: ch. 3; Mesoudi 2011: chs 3–5). The

overall order of actions is adaptive because there are systematic selection pressures that favor rule sets that promote overall orders that are more adept at facilitating cooperation and securing its social benefits, satisfying the interests and commitments of its populating agents.

7.4.2 Evaluating the Upshot of Macro Selection

This macro-selection theory of the evolution of culture is widely endorsed today. One of its implications is that we seldom can fully understand the point of evolved moral rules or the work they do for us. In his recent wide-ranging overview of cultural evolution, Henrich observes that individuals “reliant on cultural adaptations often have little or no understanding of how or why they work, or even that they are ‘doing’ anything adaptive” (2016: 57). In particular, social norms “make it possible for humans to solve—often without anyone understanding how—what would otherwise be inescapable social dilemmas” (*ibid.*: 145).

In a fundamental sense, then, the rules of social morality cannot be understood as a means to obtain social ends, for we do not know why they were selected or what social ends they serve. Hayek thus describes these rules as “purpose-independent” (1973: 81). This would certainly preclude standard consequentialist justifications of rules: we often do not know what rules do, for their effects may be linked in surprising ways to other rules and values. More generally, any systematic attempt at justifying the entire system of morality would seem doomed, for we do not know what the consequences of our entire system are, nor can we predict the effects of radical changes in a tightly coupled system of rules (Hayek 1988: ch. 5; see also Gaus 2018c). Yet Hayek (1988: 27) draws back from an evolutionary ethics in which the outcome of evolution simply defines the good. The picture is not conservative in the sense of accepting whatever is—it is, rather, of a “Whiggish” inclination (Hayek 1960: 409). Edmund Burke, a Whig admired by Hayek, was a moral critic of English policy towards the American colonies and the slave trade, while adamantly opposed to the radical social and moral reconstruction attempted by the French revolution.¹¹ From within a moral system a Whig can strenuously criticize rules, laws, and policies that violate its traditions, stepping back from this or that matter and evaluating it “piecemeal” in light of the whole (Hayek 1988: 8). But a Whig cannot sensibly step back from the whole and seek to reconstruct the evolved social morality in the light of some philosophical commitment such as utilitarianism or welfarism. “We must always work inside a framework of values and institutions which is not of our own making” (Hayek 1960: 63).

Hayek, though, sometimes suggests a more radical view. He chides the conservative for accepting whatever outcome has been produced by the latest intervention (1960: 397ff), and so the conservative’s failure to stand up for the core principles of a free order, such as liberty and property. Now interestingly, he believes that these principles can be

¹¹ See Burke (1999, vol. 1: 221–89, as well as vols 2 and 4).

derived from—or at least confirmed by—an understanding of complex systems. “A commitment to principles presupposes an understanding of the general forces by which the efforts of society are co-ordinated, but it is such a theory of society and especially of the economic mechanism that conservatism conspicuously lacks” (1960: 401). One way of interpreting this idea in a way that is consistent with a strong role for macro selection (society as a “CAS1” system) involves a sort of reverse engineering of the evolved complex system. We take an evolved system, and seek to model the ways that it came about: if we are able to do so, then we have confidence about the organizing principles and dynamics of our system (Green 2015). Armed with this theoretical knowledge, the social philosopher would be in a position to take a more radical stance than the refinement of the evolved tradition. As Hayek urges, she can stand up for the core organizing principles of the system. This, however, is a demanding research program. Reverse engineering of evolved highly complex systems is no mean feat, and serious work on it has only just begun. The question is: Do agents contemplating significant moral reform really need to pursue this demanding research agenda?

7.5 COMPLEX SYSTEMS: AGENT-BASED ADAPTATION

7.5.1 When Macro Selection Pressures Ease

D.S. Wilson (2016) is clearly skeptical of the ability of a complex system to be adaptive (that is, functional) in the absence of significant—probably quite strong—selection pressures at the system-wide level.¹² Absent macro selection, he argues, functionality is endangered. In such cases “[b]etween-group selection, if it takes place at all, is not sufficiently strong to oppose within-group selection. The groups remain functionally impaired indefinitely and there is no invisible hand to save the day” (2016: 38). Substantial evidence indicates that during the late Pleistocene era, group-level selection pressures were indeed intense. Climate was rapidly changing and groups were often thrown into severe conflict for resources (Bowles and Gintis 2011: ch. 6; Richerson and Boyd 2005: 224–9; Richerson and Boyd 2013). And in small-scale societies group competition can remain quite intense (Richerson and Boyd 2005: 206–9). Some advocates of social evolution suppose that macro selection pressures remain strong today among our large-scale societies, but this is hardly obvious. To be sure, competition exists between large-scale social orders; learning, copying, and migration continue; and in some cases this can be a source of selecting more cooperatively functional rules. But for the same reasons that group selection pressures were so strong in the late Pleistocene era—harsh, quickly changing climate, and numerous socially distinct groups near the margin of viability—we

¹² It is important that this macro selection does not entirely swamp individual-level selection; otherwise, as E.O. Wilson (2012: 243) notes, human society would look like insect colonies.

would expect that our modern era, characterized by the absence of these features, would have greatly mitigated macro-selection pressures.

But if so, it would seem that while the evolution of our cooperative and moralistic sentiments can be readily explained by a group-selection hypothesis,¹³ it is much less obvious that macro competition explains current functionality of our large-scale orders. Of course it could be that they are really dysfunctional, but compared to the hunter-gather bands of the late Pleistocene era, this supposition does not seem especially plausible.

7.5.2 Agent-Based Adaptive Complex Systems

This leads us to the second type of complex adaptive system identified by D.S. Wilson (his “CAS2” system): adaptation via adjustment by each actor to the actions of the rest. He believes that such systems could only be functionally organized by chance because of what might be called his evolutionary mindset (2016: 40): at each level, actors are seeking to maximize their fitness in competition with others, and so unless this competition is suppressed by a higher-level selection, there is no reason to think that group functional cooperation will emerge—it would be only a random event. Here economics and game theory can provide a wider framework: when the possibilities of mutual gains to agents through cooperation generally outweigh for each the benefits of individual defection, cooperative arrangements can arise simply through mutual adjustments by each actor. Consider Hume’s example of two rowers in a boat, analyzed by Vanderschraaf (2019: chs 1–2). Two men are in a boat, with each aiming to get to the other side of the river, but it takes two men to row together to do so. They are in a game of perfect coordination: no one has any incentive to defect on the functional outcome, yet there is no obvious sense in which there is group-level selection.¹⁴ As agent-based simulations such as those pioneered by Axelrod (1997), Skyrms (1996, 2004), Alexander (2007), and Vanderschraaf (2019) demonstrate, in a wide variety of contexts with different underlying games, cooperation and norms of fairness can emerge from self-organized action without any macro selection pressures.

Agent-based game theoretic models of adaptation do not typically employ robust assumptions about individual rationality and maximization; many of the formal models mentioned in the last paragraph employ simple learning rules. The great contribution of the work of Richerson and Boyd (2005: chs 3, 6) has been to show how modes of learning such as conformity bias, prestige bias, and copying successful neighbors can promote the spread of group beneficial norms in a population. We are alert to how well our

¹³ See Bowles and Gintis (2011) and Richerson and Boyd (2005: 229–31). But see Boehm (2012: ch. 7) for an argument that other forms of selection were also important.

¹⁴ A theory can always define any cooperative arrangement between two actors as a “group,” thus preserving the group selection hypothesis (Sterelny 1996). I set aside this problem: the critical question is whether behaviors can be selected that, in some context, are not in equilibrium for individuals seeking simply to pursue their own aims, such as altruism (Bowles and Gintis 2011: chs 2, 3).

neighbors are doing, for example, and as we observe them thriving we are apt to copy their actions. Because culture has evolved complex adaptive practices, humans typically do well by imitating the behavior of others around them. We often do not understand precisely the overall benefits of our cultural practices, but because culture is largely transmitted via imitation, people often do not have to know why something is done, only that it is the useful done thing around here (or the most successful). Whereas intelligent primates such as chimps tend to figure out problems for themselves, human infants have a much stronger tendency to simply copy what they observe being done, copying “stupid” acts which the chimp sees as pointless (Horner and Whitten 2005). But by copying so much, we learn a great deal from others.

7.5.3 How Can “CAS2” Systems Remain Functional?

We can identify three fundamental features of human cooperation that allow for functional “CAS2” systems. All three of these features, it is plausible to suppose, evolved in our earlier history, when there were indeed very strong selective macro pressures for group cooperation.

(i) Humans have developed modes of social organization that are hyper-social, and so most of the resources and advantages that an individual might gain are the fruits of participating in social cooperation. To put the point in less elevated terms, we are inherently tribal creatures who gain primarily through participation in groups (E.O. Wilson 2012: ch. 7). The traditional philosophical tale explaining how atomistic individuals in competition came to leave a state of nature for society is precisely that—a tale that may have some modeling uses, but has tended to create a puzzle where none exists. Humans have long since passed the point where they were in the position to attain significant advantages as solitary agents. As Henrich (2016: ch. 17) stresses, we have evolved within culture and are, at our core, cultural creatures: gaining what we need and desire through social cooperation is how we live.

(ii) To be sure, this inherently social source of benefits provides space for opportunism, in which individual advantage can be obtained through defecting on cooperation. Solving the problem of two men in a boat is easy: things suddenly become much more difficult when there are three in the boat but only two have to row (Vanderschraaf 2019: ch. 1). However, the critical Hayekian claim (discussed in section 2) is relevant here: a critical legacy of our evolutionary history is that we are as much rule-following as purpose-seeking creatures, and few are on the constant lookout to defect on rules of cooperation. As Henrich notes,

Our social psychology appears designed for navigating a world with social rules and reputations, where learning and complying with these rules is paramount... We internalize costly norms as goals in themselves, usually via cultural learning, and are particularly good at spotting norm violators, even when those violations have nothing to do with cooperation. (2016: 315–6)

(iii) As Henrich observes in the last sentence, we are experts at spotting rule violation: by four years old we are excellent cheater detectors (Cummins 1996a, 1996b). Moreover, in recent years evidence has accumulated that many are disposed to spend their own resources to punish detected cheaters, without any prospect that the punisher will be a future beneficiary from such action (Gaus 2011: 103–12). This greatly stabilizes cooperation.

Thus, by (i) we benefit mainly through social cooperation, by (ii) we are disposed to follow the social rules of cooperation, and by (iii) many of us are on the lookout for those who do succumb to opportunistic cheating and are prepared to expend resources to punish them. Given these pillars of human cooperative life, societies can maintain functional systems of cooperation even given very modest macro selection pressures.

7.5.4 Self-organization and the Return of the Invisible Hand

Those such as D.S. Wilson (2016; see also Gowdy et al. 2016: 331) tend to be dismissive of the idea of the invisible hand unless it occurs in the context of a system with fairly strong selection by group-level evolution. However, models that understand evolution simply in terms of learning rules that lead agents to adopt more advantageous cooperative strategies and norms (“CAS2” systems) seem well described as “the invisible hand of evolution” (Alexander 2007: 18). Indeed, complex systems ordered by agent-based adaptations are self-organizing in a more interesting sense than D.S. Wilson’s “CAS1,” macro-selected systems. The latter are, to a significant extent, precisely what their name implies: macro-selected. The rules under which strongly group-selected orders operate are selected via competition with other societies. To be sure, individual agents are the vehicles for instituting these rules, and they adapt their actions to them, but in a fundamental sense it was the group selection mechanism that “chose” what rules were adaptive. In contrast, in agent-based adaptive systems (“CAS2”), each agent adjusts his actions to those of others in order to better secure his ends within a system of rules. Like Smith’s invisible hand, the constituent individuals organize the order through their actions and based on their decisions: it is their mutual adjustments that gives rise to a cooperative order. In an important sense, as Hayek (1973: 29) says, such complex adaptive systems are truly the result of human action and planning but not of collective design.

In my view, Hayek, like D.S. Wilson, places too much importance on macro selection in his account of current system functionality. The result is, as we saw, an inclination to a Whiggish, Burkean approach toward altering the rules of cooperation. They are macro-functionally adaptive for reasons that must remain obscure to us (unless we achieve robust reverse engineering), so we best be careful with undertaking significant changes. The social reformer is a little like a doctor: if one knows enough and is careful, one can reliably fix modest failures of functionality, and sometimes bigger ones,¹⁵ but it is very easy to make matters worse. Contrast this to a system that is closer to the “CAS2” model,

¹⁵ The human body is far less complex than the economy; see Saari (1995).

organized primarily through individual agent-based adoptions to the actions of others. Such systems are based on constant individual feedback and adjustments. Individuals reevaluate their actions within the rules, making micro-changes on the basis of their local knowledge, and their evaluation of whether an adjustment would better promote their plans. To be sure, individuals cannot know the overall social consequences of rules, but they can decide whether the rules make their individual attempts at coordination smoother or more difficult. As Hayek approvingly notes, individuals continually test the rules to determine whether they are still serving to coordinate their plans with others:

[I]t is, in fact, desirable that rules should be observed only in most instances and that the individual should be able to transgress them when it seems to him worthwhile to incur the odium which this will cause.... It is this flexibility of voluntary rules which in the field of morals makes gradual evolution and spontaneous growth possible, which allows further experience to lead to modifications and improvements.

(1960: 63)

This individual exploration and testing makes perfect sense if the rules evolve as adjustments to agents searching for better local ways of cooperating. The rules came about through individual adaptive actions, and so continued decisions whether they remain acceptable makes perfect sense (as it makes perfect sense for others to regularly enforce them if the rules strike them as worthwhile). None of this is obviously well grounded if we think of the system as strongly macro-selected, for in a “CAS1” system the rules were not adapted because individuals found them fruitful.

Of course, any actual system of cooperation will be the result of both macro- and agent-based adaptation; I do not wish to suggest a stark choice. Yet the point remains: the more we are convinced that system functionality requires strong macro-selection, the less confident we will be that when individuals find the rules ill-suited to their individual plans, they should search for better ones and challenge the current ones. Indeed, on a multilevel selection account, the higher-level selection inherently restrains lower-level (individual) selection. *We might say, in a rough and ready way, that frustrating individual plans in order to secure system-wide functionality is precisely what macro-selection does.* “When mechanisms evolve that suppress the potential for disruptive self-serving behaviors with groups, the group evolves to be... a veritable superorganism” (Gowdy et al. 2016: 336). Given this, it is very hard to see how macro-focused accounts can accord primacy to individuals searching for “self-serving” ways of acting—those that better secure their individual aims and values.

7.6 CONCLUSION

Since Plato, political philosophy has assumed that justice or social morality is a resultant property of a set of institutions. We inquire into what sort of society is just, good, or virtuous, and then propose an institutional scheme to achieve it. Or, we select a set of

institutions and accurately predict the resulting justice of the society they produce—at least if we assume full compliance. Traditionally, that indeed has been the subject of political philosophy, and the last fifty years has been no exception. The emerging complexity analysis in economics and philosophy, however, shows this to be an illusion.

On this, macro- and agent-based adaptive accounts concur: a social philosophy that treats society and its morality as a plannable order is a fantasy. We need to take seriously “the imperative of complexity” (Page 2018). The debate between them concerns the preferred explanation of why our complex cooperative systems maintain functionality. I have argued that the more we see a complex moral order as agent-based adaptive, the more accurate it is to see the order as self-organizing: it is constantly reshaping itself, as individuals make changes, observe feedbacks, and readjust their actions to secure better cooperative or moral results. I have argued elsewhere that morality can be self-organizing in this sense (Gaus 2018d). In a similar way, Bicchieri (2016: ch. 5) models the way in which “trendsetting” individuals can commence a process of norm change that will be taken up by others in their community, resulting in a revised, self-organized, rule network. Bottom-up social self-organization is a powerful and important moral process.

To be sure, there are limits to individual adaptation and change. Social rules can be both manifestly harmful to the lives and plans of agents, yet difficult to dislodge, and so collective action is often required to eliminate or replace them. This leads to important and difficult problems: once we recognize that we are dealing with a complex adaptive system that cannot be designed in the way the democratic public would choose, we need to rethink the appropriate aims of methods of public policy. (For a general proposal, see Colander and Kupers 2014.) Hayek provided critical key insights, but serious contemporary work here has barely begun. Furthermore, we are also led to rethink political philosophy: it no longer can be the designing of an ideal set of institutions for collective choice in the just society (D’Agostino 2018; Gaus 2016: ch. 4). As Hayek’s path-breaking analysis suggested, once we begin to view social morality as a complex adaptive system, everything changes for moral, social, and political philosophy.

ACKNOWLEDGMENTS

I have greatly benefited from exchanges with my good friend Fred D’Agostino about the issues raised in this chapter. My thanks also to Jenann Ismael, with whom I conducted a graduate seminar on complexity and self-organization at the University of Arizona. I learned heaps from her, as I did from all the seminar participants, so thanks to one and all.

REFERENCES

- Alexander, J. McKenzie. 2007. *The Structural Evolution of Morality*. Cambridge: Cambridge University Press.
- Arthur, W. Brian. 2015. *Complexity and the Economy*. Oxford: Oxford University Press.
- Auyang, Sunny Y. 1998. *Foundations of Complex-systems Theories in Economics, Evolutionary Biology and Statistical Physics*. Cambridge: Cambridge University Press.

- Axelrod, Robert. 1997. *The Complexity of Cooperation: Agent-based Models of Competition and Collaboration*. Princeton, NJ: Princeton University Press.
- Bicchieri, Cristina. 2016. *Norms in the Wild*. Oxford: Oxford University Press.
- Bicchieri, Cristina, and Peter McNally. 2018. "Schemata, Scripts, and Social Norms: How Change Occurs." *Social Philosophy & Policy* 35: in press.
- Binmore, Ken. 1994. *Game Theory and the Social Contract: Playing Fair*. Cambridge, MA: MIT Press.
- Binmore, Ken. 1998. *Game Theory and the Social Contract: Just Playing*. Cambridge, MA: MIT Press.
- Boehm, Christopher. 2012. *Moral Origins: The Evolution of Virtue, Altruism and Shame*. New York: Basic Books.
- Bosanquet, Bernard. 1890. "The Antithesis Between Individualism and Socialism Philosophically Considered." In Gerald Gaus and William Sweet (eds.), *The Philosophical Theory of the State and Related Essays* (South Bend, IN: St. Augustine's Press, 2001), pp. 224–46.
- Bowles, Samuel, and Herbert Gintis. 2011. *A Cooperative Species: Human Reciprocity and Its Evolution*. Princeton, NJ: Princeton University Press.
- Brandt, Richard B. 1954. *Hopi Ethics: A Theoretical Analysis*. Chicago, IL: University of Chicago Press.
- Burke, Edmund. 1999. *Select Writings of Edmund Burke*. 4 vols. Indianapolis, IN: Liberty Fund.
- Colander, David, and Roland Kupers. 2014. *Complexity and the Art of Public Policy: Solving Society's Problems from the Bottom Up*. Princeton, NJ: Princeton University Press.
- Cummins, Denise Dellarosa. 1996a. "Evidence for the Innateness of Deontic Reasoning." *Mind & Language* 11: 160–90.
- Cummins, Denise Dellarosa. 1996b. "Evidence of Deontic Reasoning in 3- and 4-Year-Olds." *Memory and Cognition* 24: 823–9.
- D'Agostino, Fred. 2009. "From the Organization to the Division of Cognitive Labor." *Politics, Philosophy and Economics* 8: 101–29.
- D'Agostino, Fred. 2010. *Naturalizing Epistemology*. New York: Palgrave.
- D'Agostino, Fred. 2018. "How Can We Do Political Philosophy?" *Cosmos + Taxis* 5: 29–37.
- Friedman, Daniel. 2008. *Morals and Markets: An Evolutionary Account of the Modern World*. New York: Routledge.
- Gärdenfors, Peter. 2004. *Conceptual Spaces*. Cambridge, MA: MIT Press.
- Gaus, Gerald. 2006. "The Evolution of Society and Mind: Hayek's System of Ideas." In Ed Feser (ed.), *The Cambridge Companion to Hayek* (Cambridge: Cambridge University Press), pp. 232–58.
- Gaus, Gerald. 2011. *The Order of Public Reason*. Cambridge: Cambridge University Press.
- Gaus, Gerald. 2013a. "Why the Conventionalist Needs the Social Contract (and Vice Versa)." *RMM (Rationality, Morality, and Markets)* 4: 71–87.
- Gaus, Gerald. 2013b. "The Evolution, Evaluation and Reform of Social Morality: A Hayekian Analysis." In David Levy and Sandra Peart (eds.), *Hayek and the Modern Economy* (New York: Palgrave Macmillan), pp. 59–88.
- Gaus, Gerald. 2016. *The Tyranny of the Ideal: Justice in a Diverse Society*. Princeton, NJ: Princeton University Press.
- Gaus, Gerald. 2017. "Mill's Normative Economics." In Christopher Macleod and Dale Miller (eds.), *The Blackwell Companion to Mill* (New York: Wiley-Blackwell), pp. 488–503.
- Gaus, Gerald. 2018a. "It Can't Be Rational Choice All the Way Down: Comprehensive Hobbesianism and the Origins of the Moral Order." In Peter J. Boettke, Virgil Henry Storr,

- and Solomon Stein Arlington (eds.), *Buchanan's Tensions: Reexamining the Political Economy and Philosophy of James M. Buchanan* (Arlington, VA: Mercatus Center).
- Gaus, Gerald 2018b. "The Priority of Social Morality." In Thomas Christiano, Ingrid Creppell and Jack Knight (eds.), *Morality, Governance, and Social Institutions: Reflections on Russell Hardin* (New York: Palgrave MacMillan), pp. 23–57.
- Gaus, Gerald 2018c. "Hayekian 'Classical' Liberalism." In Jason Brennan, Bas van der Vossen, and David Schmidtz (eds.), *The Routledge Handbook of Libertarianism* (New York: Routledge), pp. 34–52.
- Gaus, Gerald. 2018d. "Self-organizing Moral Systems: Beyond Social Contract Theory," *Politics, Philosophy and Economics* 17: 119–147.
- Gaus, Gerald, and Shaun Nichols. 2017. "Moral Learning in the Open Society: The Theory and Practice of Natural Liberty." *Social Philosophy & Policy* 34: 79–101.
- Gowdy, John, Mariana Mazzucato, Jeroen C.J.M. van den Bergh, Sander E. van der Leeuw, and David S. Wilson. 2016. "Shaping the Evolution of Complex Societies." In David S. Wilson and Alan Kirman (eds.), *Complexity and Evolution: Toward a New Synthesis for Economics* (Cambridge, MA: MIT Press), pp. 327–50.
- Green, Sara. 2015. "Can Biological Complexity be Reverse Engineered?" *Studies in History and Philosophy of Biological and Biomedical Sciences* 53: 73–83.
- Hardin, Russell. 1999. "From Bodo Ethics to Distributive Justice." *Ethical Theory and Moral Practice* 2: 399–413.
- Hardin, Russell. 2007. *David Hume: Moral and Political Theorist*. Oxford: Oxford University Press.
- Hardin, Russell. 2013. "The Priority of Social Order." *Rationality and Society* 25: 407–21.
- Hausman, Daniel M., and Michael S. McPherson. 2006. *Economic Analysis, Moral Philosophy and Public Policy*, 2nd edn. Cambridge: Cambridge University Press.
- Hayek, F.A. 1945. "The Use of Knowledge in Society." *American Economic Review* 35: 519–30.
- Hayek, F.A. 1952. *The Sensory Order: An Inquiry into the Foundations of Theoretical Psychology*. Chicago, IL: University of Chicago Press.
- Hayek, F.A. 1955. "The Political Ideal of the Rule of Law." In *The Market and Other Orders*, Bruce Caldwell (ed.) (Chicago, IL: University of Chicago Press, 2004), pp. 118–94.
- Hayek, F.A. 1960. *The Constitution of Liberty*. London: Routledge.
- Hayek, F.A. 1967a. "Degrees of Explanation." In *Studies in Philosophy, Politics, and Economics* (Chicago, IL: University of Chicago Press), pp. 3–21.
- Hayek, F.A. 1967b. "The Theory of Complex Phenomena." In *Studies in Philosophy, Politics, and Economics* (Chicago, IL: University of Chicago Press), pp. 22–42.
- Hayek, F.A. 1967c. "Notes on the Evolution of Systems of Rules of Conduct." In *Studies in Philosophy, Politics, and Economics* (Chicago, IL: University of Chicago Press), pp. 66–81.
- Hayek, F.A. 1967d. "Kinds of Rationalism." In *Studies in Philosophy, Politics, and Economics* (Chicago, IL: University of Chicago Press), pp. 82–105.
- Hayek, F.A. 1973. *Rules and Order*. Chicago, IL: University of Chicago Press.
- Hayek, F.A. 1976. *The Mirage of Social Justice*. Chicago, IL: University of Chicago Press.
- Hayek, F.A. 1978. "Dr. Bernard Mandeville." In *New Studies in Politics, Economics and the History of Ideas* (London: Routledge), pp. 249–66.
- Hayek, F.A. 1988. *The Fatal Conceit: The Errors of Socialism*. W.W. Bartley III (ed.). Chicago, IL: University of Chicago Press.
- Henrich, Joseph. 2016. *The Secret of Our Success: How Culture Is Driving Human Evolution, Domesticating Our Species, and Making Us Smarter*. Princeton, NJ: Princeton University Press.

- Holland, John H. 2014. *Complexity*. Oxford: Oxford University Press.
- Horner, Victoria, and Andrew Whiten. 2005. "Causal Knowledge and Imitation/Emulation in Chimpanzees (*Pan Troglodytes*) and Children (*Homo Sapiens*)." *Animal Cognition* 8: 164–81.
- Kauffman, Stuart A. 1993. *The Origins of Order*. New York: Oxford University Press.
- Kitcher, Philip. 2011. *The Ethical Project*. Cambridge, MA: Harvard University Press.
- Lane, Ruth. 2017. *The Complexity of Self Government*. Cambridge: Cambridge University Press.
- Lewis, Paul. 2015. "An Anaytical Core for Sociology: A Complex, Hayekian Analysis." *Review of Behavioral Economics* 2: 123–46.
- Lewis, Paul. 2017. "The Ostroms and Hayek as Theorists of Complex Adaptive Systems: Commonality and Complementarity." In Paul Dragos Aligica, Paul Lewis and Virgil Henry Storr (eds.), *The Austrian and Bloomington Schools of Political Economy* (Bingley, UK: Emerald Publishing), pp. 35–55.
- Malthus, Thomas Robert. 1798. *Malthus—Population: The First Essay (An Essay on the Principle of Population, as it affects the future Improvement of Society, with Remarks on the Speculations of Mr. Godwin, M. Condorcet, and Other Writers)*. Ann Arbor, MI: Ann Arbor Paperback (1959 edition).
- Mesoudi, Alex. 2011. *Cultural Evolution*. Chicago, IL: University of Chicago Press.
- Mill, John Stuart. 1843. *A System of Logic*. In *The Collected Works of John Stuart Mill*, vol. 7, J. M. Robson (ed.). Indianapolis, IN: Liberty Fund (2006).
- Miller, David. 1989. "The Fatalistic Conceit." *Critical Review* 3: 310–23.
- Miller, John H., and Scott E. Page. 2007. *Complex Adaptive Systems*. Princeton, NJ: Princeton University Press.
- Mitchell, Melanie. 2009. *Complexity*. Oxford: Oxford University Press.
- O'Brien, D.P. 2004. *The Classical Economists Revisited*. Princeton, NJ: Princeton University Press.
- Okasha, Samir. 2006. *Evolution and Levels of Selection*. Oxford: Oxford University Press.
- Ostrom, Elinor. 1986. "An Agenda for the Study of Institutions." In Fillippo Sabetti and Paul Dragos Aligica (eds.), *Choice, Rules and Collective Action* (Essex, UK: ECPR Press, 2014), pp. 97–119.
- Page, Scott E. 2011. *Diversity and Complexity*. Princeton, NJ: Princeton University Press.
- Page, Scott E. 2018. "The Imperative of Complexity." *Cosmos + Taxis* 5: 4–12.
- Paul, Ellen Frankel. 1988. "Liberalism, Unintended Orders and Evolutionism." *Political Studies* 36: 251–72.
- Rawls, John. 1999. *A Theory of Justice*. Rev. ed. Cambridge, MA: Belknap Press.
- Robinson, Paul. 2000. "Why Does the Criminal Law Care What the Layperson Thinks Is Just?" *Virginia Law Review* 86: 1839–69.
- Richerson, Peter J., and Robert Boyd. 2005. *Not by Genes Alone: How Culture Transformed Human Evolution*. Chicago, IL: University of Chicago Press.
- Richerson, Peter J., and Robert Boyd. 2013. "Rethinking Paleoanthropology: A World Queerer than We Supposed." In Gary Hatfield and Holly Pittman (eds.), *Evolution of Mind* (Philadelphia: Pennsylvania Museum Conference Series), pp. 263–302.
- Robbins, Lionel. 1952. *The Theory of Economic Policy in Classical Political Economy*. London: Macmillan.
- Robbins, Lionel. 1976. *Political Economy, Past and Present*. London: Macmillan.
- Robbins, Lionel. 1998. *A History of Economic Thought: The LSE Lectures*. Steven G. Medema and Warren J. Samuels (eds.). Princeton, NJ: Princeton University Press.

- Rose, David C. 2011. *The Moral Foundations of Economic Behavior*. New York: Oxford University Press.
- Saari, Donald. 1995. "Mathematical Complexity of Simple Economics." *Notices of the AMA* 42: 222–31.
- Sen, Amartya. 2017. *Collective Choice and Social Welfare*. Exp. ed. London: Penguin.
- Simon, Herbert. 1996. *The Science of the Artificial*. 3rd ed. Cambridge, MA: MIT Press.
- Skyrms, Brian. 1996. *The Evolution of the Social Contract*. Cambridge: Cambridge University Press.
- Skyrms, Brian. 2004. *The Stag Hunt and the Evolution of Culture*. Cambridge: Cambridge University Press.
- Smith, Adam 1759. *The Theory of Moral Sentiments*. D.D. Raphael and A.L. Macfie (eds.). Indianapolis: Liberty Fund (1982 edition).
- Smith, Adam. 1776. *An Inquiry into the Nature and Causes of the Wealth of Nations*. R. H. Campbell and A. S. Skinner (eds.). Indianapolis, IN: Liberty Fund (1981 edition).
- Sterelny, Kim. 1996. "The Return of the Group." *Philosophy of Science* 63: 562–84.
- Stuntz, William J. 2000. "Self-Defeating Crimes," *Virginia Law Review* 86: 1871–99.
- Sugden, Robert. 1986. *The Economics of Rights, Co-operation and Welfare*. Oxford: Blackwell.
- Turnbull, Colin M. 1963. *The Forest People*. New York: Simon and Schuster.
- Vanderschraaf, Peter. 2019. *Strategic Justice*. New York: Oxford University Press.
- Vaughn, Karen. 1999. "Hayek's Theory of the Market Order as an Instance of the Theory of Complex, Adaptive Systems." *Journal de Economistes et des Etudes Humaines* 9: 241–56.
- Weisberg, Michael, and Ryan Muldoon. 2009. "Epistemic Landscapes and the Division of Cognitive Labor." *Philosophy of Science* 76: 225–52.
- Wilson, David S. 2016. "Two Meanings of Complex Adaptive Systems." In David S. Wilson and Alan Kirman (eds.), *Complexity and Evolution: Toward a New Synthesis for Economics* (Cambridge, MA: MIT Press), pp. 31–46.
- Wilson, Edward O. 2012. *The Social Conquest of the Earth*. New York: Liveright.

CHAPTER 8

ON THE EVOLUTION OF ETHICS, RATIONALITY, AND ECONOMIC BEHAVIOR

DAVID C. ROSE

8.1 INTRODUCTION

THIS chapter melds several strands of literature that pertain to how ethics and rationality affect economic behavior. It proposes a new metaphor, that of a bookshelf that is consulted when new circumstances arrive, which alludes to the classical approach to modeling decision-making. Although not well known by economists today, before the rise of neo-classical economics those writing on economics viewed decision-making as being driven primarily by a desire to conform to established norms.

The bookshelf does not show how decision-making has evolved from being irrational to rational; it has been rational all along. Instead, it provides a way to consider how increasing group size and evolving moral beliefs might lead to rationality being channeled so behavior increasingly comports with neoclassical economics, what McCloskey (2006) calls *Max-U*. This does not mean that *Max-U* did not exist until recently; it has existed all along. It simply became more relevant with the evolution of market economies and certain kinds of ethics. This exercise allows us to more deeply appreciate the connection between ethics and economics, and has interesting implications for the rise of individualism.

8.2 RATIONALITY

Since the ancient Greeks it has been understood that rationality is an important part of building a good individual and a good society. But while rationality helps us understand

the world and make better decisions, it can also help us act as shrewd opportunists. So even now rationality continues to be a topic of great interest to ethicists and economists, indeed anyone with a strong interest in complex animal behavior.

Of particular concern is how rationality works in group contexts. One of the most acclaimed cognitive neuroscientists of our time, Joaquin Fuster, contends that “rationality is fully compatible with group behavior if in the *perception-action cycle* of the group you include two essential attributes of human behavior that derive from evolution and are genetically transmitted: *trust* and *affiliation*.”¹ I submit that the rise of civilization has had much to do with humans improving their ability to benefit from rationality while keeping the harm it can do to the common good in check. The trick has been to extend trust beyond the reach of our small group genes through moral beliefs that make untrustworthy behavior irrational even in large group contexts.

An important lesson of economics and game theory is that rational outcomes for any society cannot be achieved by “social rationality” because social rationality is not a meaningful concept. Binmore (2005) provides an excellent discussion of this point. In short, social rationality amounts to a category error since societies do not choose—individuals do.² But even if the most rational outcome for society cannot be defined, it is still meaningful to say a society has achieved a rational outcome in that it has avoided all demonstrably irrational ones according to the Pareto criteria. But because only individuals make choices, as a practical matter even the set of Pareto efficient outcomes can only be achieved by aligning what is in the best interest of the individual with what is in the best interest of society.

Free market systems are dominated by the positive sum activity we call cooperation. Put simply, alone *A* makes 10 and *B* makes 10 but cooperating they make more than 20. There are many reasons for this outcome but one thing is always the same: the value of the whole is greater than the value of the sum of the parts. Adam Smith (1776) argued that cooperating to effectuate gains from specialization is a particularly important part of the story of how free market systems produce the good life. He argued further that the larger the group, the greater these gains will be.

The power of Smith’s argument is evidenced by the fact that large group cooperation is very common in nature. There are many species of social insects and they dominate the planet by count and mass. But their form of large group cooperation arises from precise genetic coding of behavior on a strict “if, then” algorithmic basis. This makes adapting to changing circumstances impossible within any given generation. Yuval Noah Harari (2015) argues that the key to the rise of human civilization was therefore not due to the rise of large group cooperation per se but to the rise of *flexible* large group cooperation.

Flexible cooperation refers to cooperation that can be adapted to immediate circumstances as needs arise. Whereas neighbors working together to build a fence can alter

¹ This quote is from an email to the author dated January 20, 2018. See also Fuster (2013). Later I will explain why genetically transmitted trust mechanisms are likely to weaken as we live and cooperate in ever larger groups.

² Wilson (2010) advances a related argument regarding social preferences and the folly of shoe-horning theory to fit prevailing models in response to empirical findings that show prevailing models fail to predict behavior accurately.

their plans on the fly, social insects are stuck with precisely preprogrammed behavior. Flexible cooperation requires conscious, deliberate decision-making. If decision-making reduces the likelihood of survival of the decision-maker or her group, traits that support such decision-making will die out. It follows that traits that support irrational decision-making will produce worse payoffs and will therefore die out, while traits that support rational decision-making will produce better payoffs and will therefore be reinforced. For this reason flexible cooperation almost certainly coevolved with rational decision-making.

This is not to suggest that our earliest ancestors were not rational but we today are. It is unlikely that any organism is persistently irrational. Many organisms cannot engage in rational decision-making to the extent that we do, but that does not make them irrational. It simply means that their mode of rationality is less refined. Mice are not irrational. At every moment of decision they promote their welfare the best they can with the capabilities they have. But it is nevertheless true that humans are, on average, more rational than mice. As primates evolved they evolved more sophisticated mechanisms for rational decision-making. Later, as humans evolved ever more sophisticated cultures and institutions, they also evolved an ever greater capacity for rational decision-making.

Unfortunately an increasing capacity for rationality increasingly opens the door to behavioral opportunism.³ Economists and game theorists have studied this problem for some time in the context of social dilemmas. Because we have cooperated in small groups for a very long time, we have evolved mechanisms—like feeling guilty when we harm others—to deal with social dilemmas that frequently arise in small group contexts.

As we consider ever larger groups, however, all social dilemmas worsen because of what economists and game theorists call the $1/n^{\text{th}}$ problem. As the number of individuals in a group, n , rises, the direct benefit of opportunism to the individual does not change while the cost to the individual and others from reduced group output falls. This drives up the net material payoff to the opportunist. It also drives down involuntary feelings of guilt that might arise from harming others with whom the individual can empathize.⁴ This should not be surprising. Traits that evolved in small groups, such as being good bookkeepers of favors and the ability to empathize, have no particular reason to work well in very large groups.

8.3 THE EVOLUTION OF ETHICS FROM CULTURAL PRACTICES

Flexible cooperation is not as common as large group cooperation, but it is certainly not limited to humans. Chimpanzees, elephants, orcas, wolves, and many other species clearly cooperate in a flexible way. But these species do not cooperate on anything like

³ Flexible thinking is not necessary for opportunism since genes that benefit the individual at the expense of the group or the species are supported in evolutionary equilibrium; see Dawkins (1976).

⁴ This is now known as the empathy problem; see Rose (2011).

the scale of social insects or modern humans. What is very rare and limited to modern humans is flexible large group cooperation.

For most of our existence we lived in small groups, so the problem of behavioral opportunism arising from flexible behavior was addressed by a mixture of genetically encoded traits that made us reluctant to harm others in our group and cultural practices that precluded opportunism.⁵ Patterns of behavior that made opportunism difficult were more likely to be repeated than those that left the door wide open. In this way culture, through cultural practices, addressed the problem of opportunism that grew with increasing group size through fairly precise patterns of behavior. In doing so, cultural practices increased productivity through gains from increased specialization made possible by reducing the risk of opportunistic exploitation.⁶

By addressing the increasing problem of larger group size leading to more opportunism, the mediation of behavior through cultural practices led to an increase in the size of human groups. But increasing the degree to which behavior follows precise patterns also has the effect of bottling up rationality. The mind automatically searches for the appropriate “then” response when confronted with any given “if” circumstance.

This has the effect of taking most or all of the remaining action set out of consideration. When circumstances never change, or the same small set of circumstances are repeated without variation, this is of no concern. But even small changes in circumstances can lead to an action that was not prescribed by the relevant cultural practice being the best action. Strict cultural practices therefore reduce the likelihood that actions that might better promote the common good will be selected. This reduces the ability of individual rationality to produce diverse decisions that can begin new evolutionary paths of decision-making in the future. To the extent that some of these paths might have increased the common good, this negates some of the social advantages of flexible cooperation.

Intergroup competition produced rewards to those groups that could cooperate in large group contexts to benefit from efficiency gains from increased scale. It also produced rewards to groups that could cooperate in a flexible rather than pre-programmed fashion. As groups evolved mechanisms that could better support large group cooperation or flexible cooperation, they came to dominate those that could not.

Both large groups and flexible decision-making open the door wider to rationality being exercised to promote individual welfare at the expense of the group. This tends to keep flexible groups small and large groups inflexible. But this also means that the advantages of flexible large group cooperation should reinforce traits that combat opportunism in large group contexts to make flexible large group cooperation possible. It follows that as groups grew larger, those that found ways to better combat large group opportunism would be able to more fully unleash the power of flexible large group cooperation and dominate those that did not.

⁵ See Boyd and Richerson (1985) and Richerson and Boyd (2005).

⁶ Thanks to the pioneering work of Robin Dunbar, it is now widely understood just how small-group oriented modern humans still are. This work explains how many institutions and organizations evolved in ways that deal with our small-group limitations (see Coward and Dunbar 2014 and Dunbar 2016).

One way to break open the floodgates of flexible large group cooperation would be for specific kinds of moral beliefs to evolve that happened to pre-rationally foreclose opportunistic actions. This would address the problem of opportunism by robbing it of the power of rational calculation while otherwise leaving rationality in play for scientific inquiry and wise decision-making. So perhaps institutions did not become more important than culture to unleash the modern world. Perhaps, instead, cultural *beliefs* became more important than cultural *practices*, and then certain cultural beliefs made possible certain kinds of institutions, such as highly trust-dependent institutions, that in turn dramatically increased average productivity.⁷

8.4 THE EVOLUTION OF ECONOMIC BEHAVIOR

Adam Smith (1759) explained how group norms, through group approval and disapproval, can emerge to provide a means of group adaptation to specific environments. Recently Vernon Smith and Bart Wilson (2017) have argued that some of the shortcomings of *Max-U* disappear when we think of behavior as being governed as described by Adam Smith. They explained why the desire to conform might provide a better explanation for the predictive failure of the expected utility hypothesis than other-regarding preferences that work through mutual affection, a taste for fairness, or predilections for positive or negative reciprocity. Significantly, Smith and Wilson stress that their analysis pertains to behavior in small groups.

Taking inspiration from both of these sources, I will explain how economic decision-making might have evolved from a mode preoccupied with conforming to established norms and practices in small groups to a mode that is better described by *Max-U* in large groups. I will later explain how this process might have been catalyzed by the evolution of ethics that channel rationality in a uniquely beneficial way for societies that seek mass flourishing.

8.4.1 The Neoclassical Emphasis on *Max-U* and the Classical Emphasis on Conformity

When a new circumstance arises, *Max-U* says the first thing a person thinks of is essentially “Given how this circumstance affects my constraints, how do I maximize my welfare?” This question is answered by the model of *rational choice* in which behavior is driven by the individual trying to maximize her utility given her personal tastes.

⁷ See Rose (2019).

In the traditional formulation what is valued by the individual are things that improve the individual's welfare irrespective of how others are affected or what others might think. Social norms, mores, folkways, taboos, and moral values are not considered. This is why the traditional formulation of *Max-U* is often equated with narrow self-interest.

But even with this incredibly simple formulation of utility the best response is often not obvious. The best response for the individual maximizes utility in light of all constraints, which requires a careful consideration of all the possible costs and benefits—the primary mode of rational analysis. Rational decision-making as conceived by *Max-U* therefore involves the executive system of the brain, specifically what might be called the cost-benefit analysis center.

When a new circumstance arises, the classical view of behavior (henceforth *Max-C* for “maximize conformity”) says the first thing the person thinks of is “Given this new circumstance, what I am supposed to do?” Adam Smith explained why in most cases the answer comes from an established response that has evolved over time in her group. Such socially proper responses are followed automatically for the most part, like an “if, then” algorithm in a computer program.

Even in large modern societies today, much of our daily behavior amounts to simply trying to follow the rules of proper conduct. We simply do what we think we are supposed to do and don't do what we think we are not supposed to do. When an unattended child scrapes her knee we bend over and help her up without even thinking about it. When someone dies we don't even have to not think about not hiring a clown to perform at the funeral.

8.4.2 A Bookshelf Metaphor

To set the stage for seeing how behavior that comports mostly with *Max-C* can evolve into behavior that comports mostly with *Max-U*, envision a large bookshelf. When a particular circumstance arises, the individual automatically consults the bookshelf to find the proper response, which is contained in a particular book. For most circumstances such behavior takes the form of an “if, then” algorithm. An extreme example of this would be the sharp “if, then” behavior of social insects for which genes write very specific behavioral responses indeed.

Responses that are so strongly disapproved by society that they are completely beyond the realm of contemplation for a given circumstance are signified by books that have a cover that is pure red. Responses that are so strongly approved by society that they are absolutely imperative are signified by books that have a cover that is pure green. Responses that are completely neutral are signified by books that have covers that are pure brown. Other responses are distributed along a continuous hue scale between pure green and pure red.

For every known circumstance the pattern of book colors can differ because different circumstances require different responses. If there is only one pure green book, the individual immediately selects it and responds as it directs. This is a case of a circumstance

that has a morally required response like throwing a lifesaver to someone who just fell overboard. For most circumstances in life there is no pure green book.

One way to think of this hue scale is that it informs what Adam Smith (1759) calls the individual's impartial spectator. The various hues tell the individual how most other persons in her group would react to her having chosen any given response. So the hue reflects the individual's understanding of how strong the approval (brown/green to pure green) or disapproval (brown/red to pure red) will be for each response. This helps the individual make good decisions by giving the individual a mechanism through which to imagine, before choosing a course of action, the likely reaction of others after making such a choice.

The impartial spectator helps provide strong conformity in a group by taking the place of others who might offer approval or disapproval when they cannot be present or before they will be present. The individual feels good about doing what would be expected to arouse approval and feels bad about doing what would be expected to arouse disapproval even if that response is unlikely to actually occur because her actions are unlikely to be observed. This is simply taking comfort in knowing you did what you were supposed to do even when no one else does, and can be viewed as a means by which behavior that is normally mediated by external shame can also come to be mediated by internalized guilt because we have a conscience (Lal 1998).

Few would quibble with Adam Smith's claim that we are very sensitive to the approval and disapproval of others. If every time a child responds in a particular way to a specific circumstance she receives strong and universal social disapproval, the book associated with that response will become increasingly red. If every time a child responds in a particular way to a specific circumstance she receives strong and universal approval, the book associated with that response will become increasingly green. If every time a child responds in a particular way to a specific circumstance she always receives neither approval nor disapproval, the book associated with that particular response will remain brown or become increasingly brown.

For every circumstance, then, the bookshelf will have a particular pattern of hues because the group expects different responses to different circumstances. The book that contains the response "go to the ballgame" is green when the circumstance is your daughter wants you to take her to the ballgame for her birthday. But that same book is red when the circumstance is a death in the family. As a child grows up, the bookshelf gets larger and the pattern of colors increasingly come to reflect the pattern that is common to the minds of all other adults in the society. The smaller and simpler a society is, and the slower its rate of change, the truer this is. So the smaller the society and the more refined are cultural practices due to a long period of stability, the stronger will be the habit of mind to think first about what one is supposed to do to.

The smaller the group, the truer it is that conformity is well defined so the easier it is to conform. At the same time, the smaller the group the more likely everyone knows everyone else, so the smaller the group the more important it is to conform. Obviously the easier and the more important it is to conform, the more likely the dominant habit of mind will be to respond to each new circumstance by first consulting the bookshelf to find the appropriate response so as to conform.

Some situations are completely novel. Such situations necessarily require careful judgment, including but not limited to rational cost-benefit analysis. But no matter how large a society becomes, there will always be some actions for which nearly every individual will have an automatic sense of what to do. On a Brooklyn beach this summer there will almost certainly be a toddler who wanders too far into the surf and a complete stranger will dash out to grab the toddler. It will almost certainly be the case that the stranger will not do a rational cost-benefit analysis. If asked, she will later say she just did what she had to do without thinking. And with language that would delight Adam Smith, she might even insist that she just did what anyone else would have done.

When there are two or more proper responses, the “if, then” nature of the bookshelf does not automatically produce a unique course of action, and the individual has to make a choice. In such cases there is no reason why the individual cannot choose the alternative that best promotes her own welfare. So when behavior largely comports with *Max-C*, individual rationality takes on a well-defined role: it breaks ties. To determine the alternative that best promotes the individual’s welfare, the individual rationally considers the costs and benefits involved which is informed by the individual’s personal tastes. Note that red books normally do not require tie breaking because one can normally refrain from any number of things simultaneously.

Since the brain is a very energy-intensive organ, evolution naturally favors modes of decision-making that avoid wasting resources. The same process by which the bookshelf is built affects how individuals make decisions by affecting how neural pathways are created and destroyed in the brain. When a particular circumstance arises a child has to think about what to do, which amounts to routing the decision of how to respond to that circumstance through the executive system to conduct cost-benefit analysis. Suppose each time a particular response is considered for a particular circumstance the answer is always no. In other words for a given individual, circumstance x_1 produces the answer “no” for response y_1 , over and over again.

A well-known principle of cognitive science is that neurons that fire together wire together.⁸ Let us now explore how this principle applies to the bookshelf. The arrival of circumstance x_1 requires the consideration of many y options for a response. Should the individual do y_1 in response to x_1 ? The younger she is, the more likely this is an open question that is routed through the cost-benefit analysis center.

Suppose after doing so, time and time again, the answer is always “no.” In this case neurons will begin to wire together the question “If x_1 , should I do y_1 ?” to the answer “no,” effectively looping around the cost-benefit analysis center. Since the answer is always the same the exercise proves superfluous. So at the same time the neurons routing the question through the cost-benefit analysis center begin to wither.

The strengthening of the neural connection between the question above and the answer “no” is analogous to the book that corresponds to the y_1 response becoming red-der over time with respect to x_1 . As a result a conscious rational consideration of whether

⁸ This is widely attributed to Löwel (1992) but it expresses the main idea of Hebb’s (1949) theory of neuroscience.

to undertake y_1 is increasingly unlikely to be undertaken. Since normally one can refrain from taking any number of negative moral actions, tie breaking is normally not required. By the time the book approaches pure red the neurons that once led this question through the cost-benefit analysis center will have long ago withered into oblivion.

Now suppose the question is “If x_1 , should I do y_2 ?” and the answer is always “yes.” In this case neurons will begin to wire together the question “If x_1 , should I do y_2 ?” to the answer “yes,” also looping around the cost-benefit analysis center. And again, since the answer is always the same the exercise proves superfluous. So the neurons routing the question through the cost-benefit analysis center begin to wither.

The strengthening of the neural connection between the question above and the answer “yes” is analogous to the book that corresponds to the y_2 response becoming greener over time with respect to x_1 . As a result, a conscious rational consideration of whether to undertake y_2 is increasingly less likely to be undertaken. But taking a given positive moral action often requires resources that cannot be used for taking other positive moral actions. This means ties must be broken because, unlike negative moral actions that can be simultaneously avoided, more than one positive moral action cannot be simultaneously taken with the same resources. In such cases equally compelling yes answers will have to be routed through the cost-benefit analysis center.

When the answer is always the same because of the consistency and strength of social approval or disapproval, as the individual ages the decision about how to respond slowly ceases to be a decision in the normal sense of the word and becomes an automatic response. This comports with daily experience. A great deal of behavior that is essentially automatic to us as adults was far from automatic for us when we were children. As adults we might now call these instinctive responses, but since they are not based on genes but on social learning over time, it is better to call them intuitive or automatic responses. In this way, an individual can come to think increasingly as described by *Max-C*, as one who conforms by striving to do or not do as required by social norms.

Responses to circumstances that are directly connected to a yes or a no are not subjected to rational cost-benefit analysis but this does not mean that they are irrational. Since the response is made ahead of rational analysis, from the perspective of the individual it is effectively pre-rational. And since such responses are made because of cultural practices that evolved at the group level, they are likely rational when considered in terms of group welfare.

Genuine trust is what we have for those whom we believe will not betray us even when there is no chance of being detected. Robert Frank (1988) called such circumstances *golden opportunities*: chances to behave in an opportunistic way when the would-be opportunist believes there is no possibility of detection. Before Frank the word trust was often equated with merely having confidence that things would work out as expected. The concept of a golden opportunity is important because it clarifies that the word trust is meant to convey the idea of moral trust rather than, for example, trust in another person's competence.

The concept of golden opportunities also helps clarify the distinction between genuine trust and what Oliver Williamson (1993) calls “calculative trust” and Toshio Yamagishi

(2000) calls assurance. The more specialized economic activity is, the more localized knowledge will be and therefore the more likely golden opportunities will arise. So the rise of genuine trust is important for keeping transaction costs low in very large and highly specialized societies because in such societies it is frequently the case that individuals will believe that they have no chance of being detected if they behave opportunistically.

Frank's point was, in part, that it is mistaken to view genuine trust as irrational even though it requires not taking advantage of golden opportunities. If individuals possess traits that make them trustworthy, so they can be trusted even if golden opportunities will likely pass their way, and if they also possess involuntary emotional responses that are reliable indicators of having such traits, then such individuals can benefit because trustworthy transaction partners are more valuable. The short-term losses from not acting on this or that opportunity can pale in comparison to the long-term gains from being believed to be trustworthy.

Moreover, the group can benefit as well and, as such, the individual can benefit indirectly by virtue of being a member of the group. That which can achieve group-level rationality is often irrational when judged solely in light of the individual's welfare at the moment of a given decision. For example, being taught that one should immediately harm those who harm the group's children will result in less harm being done to the group's children. This benefits the group, but it may end up getting the retaliator killed. When judged solely from the perspective of the retaliator, beliefs, practices, or emotions that produce retaliation may be irrational given the risk involved, but the average payoffs for individuals in the group may be higher because of the benefits to the group as a whole.

What is optimal and therefore rational for the group, then, might not be rational when judged solely from the individual's perspective. So when decision-making comports with *Max-C*, group-level rationality can produce outcomes that are contrary to individual rationality at the moment of decision. But that does not mean that such an individual would prefer living in a society where group rationality did not change the outcome from what behavior driven solely by individual rationality would produce.

Now consider a circumstance for which the bookshelf offers little guidance (all books are better described as some kind of brown than green or red). For example, you are 12 years old and you are asked, "Would you like to see Grandma at the hospital?" You know you are supposed to say yes, of course. But then you are asked, "Should we bring her a card or some candy?"

In this second case it is not clear what the right answer is since they both appear likely to elicit social approval so they both amount to answers contained in equally green books. So you have a tie to break. Which do you pick? You rationally choose what you think will make *you* happiest by rationally comparing the costs and benefits to you. Perhaps after a few seconds you conjecture that Grandma might share some of her candy but you really don't want to eat part of her card, so you pick candy.

Long ago in very small group life, cultural practices told us what to do in many and perhaps most circumstances. But in the modern world we often have to make decisions for which established cultural practices do not tell us precisely what to do. Because we are empathetic by nature, in such cases if others' welfare might be involved we would

naturally try to imagine how our decision might affect them, especially those we care much about, and this often guides us to a decision. But in the modern world such considerations are often irrelevant. In such cases one is free to simply choose what one prefers given one's personal tastes without fear of disapproval or even fear of arousing one's own feelings of guilt. So in such cases there is no reason not to break ties by simply asking yourself, "Which of the options is best for me?"

Because the answer to that kind of question might not be obvious, such questions are often routed through the executive system where a careful, rational, cost-benefit analysis is applied. Breaking ties therefore involves behavior that looks less like *Max-C* and more like *Max-U*. Since in most cases the costs or benefits are affected by the individual's personal tastes, the response ends up being strongly affected by the individual's tastes.

8.5 THE EFFECT OF GROUP SIZE ON ECONOMIC BEHAVIOR

With increasing group size, specialization increases, so the number of possible circumstances rises. This makes it harder to identify, instantiate, and reinforce specific responses. The number of circumstances for which established cultural practices provide clear direction falls relative to the number of circumstances for which they do not.

Increasingly new circumstances arise that have too little in common with known circumstances to provide a basis for analogous reasoning. For example, your child wants to download a game app on her smartphone. Is this proper? If you're of a certain age, no question even remotely similar arose in your own childhood, so your parents' behavior offers no guidance. At the same time, cultural practices are not as strong—the books are rarely close to pure green or pure red—because things change so quickly that consistent patterns of approval and disapproval do not have enough time to be clearly discerned. The bookshelf-building process has trouble keeping up.

Increasingly, therefore, individuals reach adulthood with a great many brown books on the bookshelf with respect to a great many circumstances, so they will have little or no idea what to do based on established norms of social propriety. At the same time, it is increasingly often the case that our capacity to empathize with possibly affected parties is moot because the group is too large for anyone to be meaningfully benefitted or harmed.

So there will be more circumstances for which the only immoral responses are responses that are always immoral because they are categorically immoral. Such books are always red, as in the case of murder. Other responses are always immoral because the group will have evolved beliefs and practices that effectively deal with the empathy problem. For example, we are taught in our group never to lie because lying is inherently wrong, so it is wrong even if no one is harmed. This defeats the $1/n^{\text{th}}$ problem and therefore the empathy problem.

With increasing group size, there will be many more circumstances in which established cultural practices tell the individual strongly what *not* to do but do not tell the individual *what* to do. Think for a moment of how many decisions you will make today that are

effectively non-moral in nature. In such cases it never even occurs to you to consider which action is the most likely to arouse moral approval or disapproval. You automatically *don't think* of a long list of immoral responses. Some are clearly immoral to you because of their inherent nature (do not murder), but some are clearly immoral to you because you have learned that they are to be regarded as categorically immoral (do not cheat, such as on your federal income taxes). Your decision-making is therefore mostly trained on breaking ties, which you do by asking yourself which response best promotes your welfare given your personal tastes. Only rarely in daily modern life do we come across circumstances that have pure green books: perhaps if you just learned your best friend's mother has died then you should call her.

It follows that as group size increases, decision-making becomes less about pulling down a scripted response for a given circumstance and becomes more about making a choice that makes the individual happiest within a number of constraints: some material, some social, and some moral. So increasingly individuals make decisions in a way that is better described by *Max-U* than *Max-C*. This puts individualized rational thinking far more in play than in the small group world in which behavior is well described by *Max-C*.

So the larger the group the harder it is to conform, and the less important it is to conform. In many more circumstances there are no books with pure green covers because there are too many people and too many novel circumstances for a clear consensus to emerge, so a specific proper response is not well defined. The individual focuses less on conforming vis-à-vis automatic obedience of cultural practices and more on what the individual wants, because what the individual wants simply matters more in larger group contexts.

Because moral proscriptions are inherently more objective than moral prescriptions—"do not murder" is well-defined while "be generous" is a matter of degree, therefore not well-defined and chosen in part because of individual tastes—group norms come to function less as precise patterns to be imitated and more as sharp and inviolate constraints within which to choose. As group rationality is imposed ever more through proscriptions than prescriptions for behavior, individual behavior comports ever better with the mathematical structure of *Max-U*.

Increasingly over time a new habit of mind emerges, which is to ask "Given this change in circumstance, how can I best promote my welfare?" This habit of mind arises because, over time, the neural path to rational analysis is reinforced because automatic responses produce worse payoffs on average than carefully considered ones. Therefore more decision-making is subjected to individual rational scrutiny. Since rational analysis involves weighing of outcomes according to personal tastes, this means decisions will increasingly vary across individuals because they will reflect the diversity of personal tastes that in small group life would have been largely suppressed by average group conceptions of moral propriety.

When most decision-making is as described by *Max-C*, the power of individual rationality is largely bottled up, often producing outcomes that are rational for the group but not for the individual. But even the group suffers when it fails to benefit from the power of individualized rationality, because individuals possess diverse tastes and diverse beliefs.

More generally, decision-making as envisioned by *Max-C* effectively forecloses a great deal of rational thinking. Precisely scripting responses to circumstances has the effect of redacting all other responses from the action set. This, in turn, severely limits the application of logic and reason to the analysis of the world and to making decisions generally. Given how resource intensive the brain is, this virtually assures atrophy of the mechanisms that support rational analysis. Better to just do what you are supposed to do and not think further.

In contrast, decision-making that comports with *Max-U* doesn't prescribe specific responses. Instead it proscribes certain responses through social and moral constraints that are added to the customary constraints in the rational choice model. This leaves a great many more socially permissible responses to any given circumstance. The diversity of personal tastes can thereby produce a diversity of responses, which produces a richer set of background conditions for future decisions. This richness reinforces itself and comports with the dramatic degree of creativity and diversity we find in large and cosmopolitan societies.

With respect to actions that involve production, so the main concern is the creation of wealth to support consumption later, opening up the action set makes it more likely that the individual will make choices that maximize the size of the cooperative or exchange surplus because this is in her own best interest. But these are precisely the actions that best promote the common good. As long as negative moral actions are not allowed—actions whose redaction constitutes additional constraints within which permitted choices are made—self-interest naturally takes the individual in directions that end up maximizing the value of output per person in society. This is the path to mass flourishing.

8.6 THE EFFECT OF ETHICS ON ECONOMIC BEHAVIOR

As intergroup competition drives groups to become ever larger, decision-making increasingly involves making conscious, deliberate, rational decisions rather than automatically doing as the bookshelf directs. Larger group size also weakens social ties, so we sympathize less with harm that might come to others from our opportunistic behavior. Also, larger group size often spreads harm over so many persons that there is no harmed individual with whom to empathize, so even basic decency arising from modest sympathy for strangers can be rendered moot.

Unfortunately the cumulative effect of such opportunism can seriously undermine the common good (Rose 2011, 2016). Increasing group size catalyzes the gains from specialization and leads to more decision-making that is rational and therefore more likely to discover efficient actions than consulting the bookshelf. But at the same time, increased specialization localizes knowledge, which in turn increases opportunities for

opportunism while often rendering moot our natural reluctance to engage in opportunism because of the absence of noticeable harm done to others.

Harari (2015) argued it was through the creation of institutions that humans were able to enjoy flexible large group cooperation because, in part, they helped us overcome the increasingly difficult problem of opportunism. But perhaps the story doesn't end there. Institutions, including those associated with government and religion, took human civilization to unprecedented levels. Many civilizations, however, were highly dependent on brutal punishments and enslavement. Until recently in history, none were capable of producing anything like mass flourishing.

I propose that incredible gains from even more flexible cooperation in even larger groups were made possible by more effective suppression of opportunism, and it was this suppression of opportunism that was the key to producing mass flourishing. Scientific achievements get most of the credit, but they cannot be the whole story. Far too many people live in societies that fall far short of mass flourishing but have access to such knowledge. They are also filled with smart and highly trained engineers and scientists, have abundant natural resources, and enjoy the benefit of richer nations ready to help with expertise and institutional templates.

What these societies lack is a high level of trust made possible by the strong suppression of opportunism. Perhaps it was a change in the nature moral beliefs that helped produce an ethic of duty-based moral restraint which, in turn, produced trustworthiness that remained intact in large group contexts. Duty-based moral restraint refers to an unwillingness to even consider taking negative moral actions. It produces a lexicographical ordering of moral preferences in that the moral value of positive moral actions is only considered if no negative moral actions are undertaken. This is particularly important for supporting large group cooperation, because without duty-based moral restraint, it is often easy to rationalize taking negative moral actions as means to taking positive moral actions, because the harm done to those with whom we can empathize and sympathize can be driven very low when divided among a large number of individuals.

Duty-based moral restraint essentially takes certain actions off the table by removing them from even being considered by the individual. This leads to an unbroken pattern of "no" answers after cost-benefit analysis has been performed. In the brain this leads to looping around the executive system and therefore denies rationality a handhold for rationalizing the taking of negative moral actions. In doing so it aligns individual rationality with group rationality and thereby best promotes the common good. It provides a basis for flexible large group cooperation that institutions cannot provide because they do not work in cases of golden opportunities. It likely also changed the path of the evolution of economic behavior by driving decision-making even farther away from *Max-C* and closer to *Max-U*.

Paradoxically, this sharp curtailment of rationality with respect to negative moral actions has the effect of strongly supporting rational thinking. Having removed the set of responses that undermine the common good, there is no need to otherwise circumvent the executive system. Rationality is therefore free to be applied to understanding the

world through scientific inquiry and to assist with wise decision-making that is within the bounds of ethical behavior. Moral beliefs that stress duty-based moral restraint therefore allow us to be more fully rational.

Duty-based moral restraint therefore allows humans to build societies with high levels of diversity of thought, rationality, and trust. Such societies can support highly trust-dependent institutions that further unlock the power of cooperation. Such societies are very much like what we envision when we envision a good society within which we flourish both individually and collectively. Such societies allow us to be diverse individuals, to fully manifest a condition of individualism that has nothing to do with selfishness. By cutting off the opportunistic path to success, duty-based moral restraint essentially forces those who want to be successful to focus on cooperating with others as effectively as possible and on the greatest scale possible.

8.7 INDIVIDUALISM

The smaller are groups and the weaker is the belief that moral restraint should take precedence over moral advocacy, the truer it is that as new circumstances arise individuals will first think of what they are supposed to do given group norms. Unless a tie needs breaking, for the most part rational decision-making begins and ends with consulting the bookshelf in an effort to conform to the group's collective view of moral and social propriety. This may help explain why individualism is far less evident in very small group societies. At the same time, it may help explain why individualism emerged so fully in places like America.

With increasing growth and development the bookshelf has a hard time keeping up, so by the time an individual reaches adulthood many circumstances produce a great many brown books and some red books. Increasingly, the bookshelf does not tell the individual what to do, but only tells the individual what not to do, so the individual's decision-making requires tie breaking. Because individuals will naturally favor responses that break ties in a way that benefits them, they get used to thinking less about conforming and more about how their choices might benefit them and those they care most about.

Therefore, when economic behavior comports with *Max-U* it produces habits of mind that comport with individualism. The first response to a new circumstance becomes to consider how best to promote one's own welfare given the effect of the new circumstance on constraints. As a result, behavior increasingly reflects the diversity of individuals through the diversity of their tastes, and individuals have more reason to think about themselves and what they want rather than what they are supposed to do as members of a group. In this way economic decisions come to increasingly reflect and contribute to a sense of individuality.

Another way to think of how small group context naturally suppresses a sense of individuality is to recognize that small group society has the paradoxical effect of diminishing the importance of the individual. Because the emergence of standards for behavior arise

at the group level and the bookshelf is most pertinent for small groups, it follows that the smaller the group, the more often that individuals behave in ways that have more to do with promoting group welfare than the desires of the individual. In many cases the unit of decision-making analysis in a small group society is the group, while in many cases the unit of decision-making analysis in large societies is the individual.

Individualism should therefore not be dismissed as selfishness. Selfish persons do not mind harming others to help themselves, but those who are individualistic because they abide by an ethic of duty-based moral restraint would never harm others as a means to the end of benefitting themselves. Philip Wicksteed's (1933) concept of *non-tuism* is a more accurate characterization of how strangers relate to one another in a large individualistic society made possible by an ethic of duty-based moral restraint. Non-tuism refers to not being terribly concerned with making those not close to us happy or with behaving in a noble way, but nevertheless obeying the rules of civilized behavior.

Non-tuism comports with the idea of thin social trust. Thick trust is like the trust a child has in her mother. It is very deep but it is rooted in mutual affection which is by nature limited to a small number of persons. Thin trust is like waiters not worrying that tips won't be left for them. It does not work for great sums of money, but because it is not derived from mutual affection but, instead, from social norms, it can apply to a great many people, even to society in general.

With thin trust in large societies, it is hard to trust randomly drawn strangers in a deep way, but it is easy to not be suspicious of randomly drawn strangers (for example, to trust them not to cheat us). This is crucial for producing a sense of easy anonymous comity whereby we are not burdened by a large concern for others but we nevertheless benefit from not having to fear that others will exploit us, either. This a condition that exists in many very large Western societies and China.⁹ Such a mindset is critical for any society that wishes to benefit from honest competition even though it can be expected to produce outcomes that harm some individuals.

8.8 CONCLUSION

Ethics that effectuate duty-based moral restraint channel rationality in a way that eliminates opportunistic actions while leaving rationality otherwise free to help individuals promote their self-interest, make the best possible decisions, and make the most sense of the world around them. This catalyzes the effect that increasing group size has on moving decision-making in a direction that is better modeled by *Max-U* than *Max-C*. This also helps free the mind for more open-ended rational inquiry into better understanding our world through the scientific method and into making better decisions both individually and collectively.

Because under *Max-U* decisions will be governed more by individual tastes, it follows that it contributes to individualism. One's identity and self-worth is therefore no longer

⁹ See Ortiz-Ospina and Roser (2017) and Rose (2018).

derived mostly from being a part of a group and therefore a means to the end of protecting and promoting a particular group's collective welfare. Instead one thinks of oneself as a unique individual whose welfare is an end in itself.¹⁰ This exercise therefore helps connect the emergence of individualism to the rise of large group modernity, *Max-U*, and to the rise of moral beliefs that instantiate an ethic of duty-based moral restraint.

Experimental economics has revealed a number of problems with *Max-U*. But how much of this is a problem with *Max-U* per se versus a problem with applying a model best suited for individual decision-making in large group contexts to behavior that normally takes place in small group contexts? Most trust games, for example, are implicitly framed in a small group context and therefore can be expected to actuate small group moral intuitions such as acute empathy even for strangers and the suspicion that repeat play is in play even if it is said that it is not.¹¹

As Smith and Wilson (2017) explained, what I have dubbed *Max-C* is perhaps a better model of behavior in small group contexts than *Max-U*, even after modification by the various accoutrements and workarounds of modern behavioral economics. No one expects an ocean liner to perform well in a river. Modifying the ocean liner might help, but the smaller the river the clearer it becomes that a better approach would be to get a smaller boat. This suggests that, given what we now know about shortcomings of *Max-U*, what is needed is a broader model for which group size is a parameter that takes behavior from that described in *Max-C* to *Max-U* as group size increases.

When n is very small, decision-making is primarily a matter of striving for conformity by doing, as precisely as possible, what is required as informed by the bookshelf. In such a world rationality exists, but it exerts its force on decision-making mostly at the group level. As n grows, the ratio of circumstances that produce ties to those that do not rises, so rational application of cost-benefit analysis comes increasingly into play. Group standards still matter, but increasingly they proscribe action more than prescribe it. As n grows even more, this ratio increases and the extent to which social constraints on decision-making is lessened further, so *individual* rationality becomes even more important and the pattern of resource allocation comes to reflect even more the rich diversity of tastes that vary by individual.

The ancient Greeks believed that rationality was intimately related to the good, both for the individual and for society as a whole. Surely they believed, or at least hoped, that there is a way to construct society so both conceptions of the good are compatible. Their work, and the work of countless scholars that followed, searched for that connection.

¹⁰ Note the historical timing of Immanuel Kant's writings in the late eighteenth century that reinforce the idea that persons are not to be used as means to others' ends, and the rise of very large societies with an increasingly evident prevailing ethic of duty-based moral restraint over a prevailing imperative to conform to behaving as expected by one's group.

¹¹ Note that Hayek (1988) argued that capitalism constituted an extended order of large group cooperation. In his view such large group cooperation was made possible by a legal framework that he viewed as providing a substitute for trust, which he viewed as a small group phenomenon.

The rise of large free-market societies and the evolution of moral beliefs that produce something like an ethic of duty-based moral restraint can be seen as two factors that helped channel rationality so as to make what is best for the individual also be that which is best for the common good.

ACKNOWLEDGMENTS

The author gratefully acknowledges support by the Earhart Foundation, the Templeton Foundation, and the International Studies and Programs center at the University of Missouri-St. Louis. The author also thanks Robin Dunbar, Robert Frank, Joaquin Fuster, James Otteson, Maria Paganelli, and Bart Wilson for comments and suggestions on an earlier draft.

REFERENCES

- Binmore, Kenneth. 2005. *Natural Justice*. New York: Oxford University Press.
- Boyd, Robert, and Peter J. Richerson. 1985. *Culture and the Evolutionary Process*. Chicago, IL: University of Chicago Press.
- Coward, Fiona, and R.I.M. Dunbar. 2014). "Communities on the Edge of Civilization." In R.I.M. Dunbar, Clive Gamble, and J.A.J. Gowlett (eds.), *Lucy to Language: The Benchmark Papers* (Oxford: Oxford University Press), pp. 380–405.
- Dawkins, Richard. *The Selfish Gene*. 1976. New York: Oxford University Press.
- Dunbar, Robin I.M. (2016). *Human Evolution*. New York: Oxford University Press.
- Frank, Robert H. 1987. "If Homo Economicus Could Choose His Own Utility Function, Would He Want One with a Conscience?" *American Economic Review* 77: 593–604.
- Frank, Robert H. 1988. *Passions Within Reason: The Strategic Role of the Emotions*. New York: W.W. Norton.
- Fuster, Joaquin. 2013. *The Neuroscience of Freedom and Creativity: Our Predictive Brain*. Cambridge: Cambridge University Press.
- Harari, Yuval Noah. 2015. *Sapiens: A Brief History of Humankind*. New York: Harper Collins Publishers.
- Hayek, Friedrich A. 1988. *The Fatal Conceit: The Errors of Socialism*. W.W. Bartley, III (ed.). Chicago, IL: University of Chicago Press.
- Hebb, Donald. 1949. *The Organization of Behavior*. New York: Wiley & Sons.
- Lal, Deepak. 1998. *Unintended Consequences: The Impact of Factor Endowments, Culture, and Politics on Long-Run Economic Performance*. Cambridge, MA: MIT Press.
- Löwel, Siegrid, and Wolf Singer. 1992. "Selection of Intrinsic Horizontal Connections in the Visual Cortex by Correlated Neuronal Activity." *Science* 225: 209–12.
- McCloskey, Deirdre N. 2006. *The Bourgeois Virtues: Ethics for an Age of Commerce*. Chicago, IL: University of Chicago Press.
- Ortiz-Ospina, Esteban, and Max Roser. 2017. "Trust." OurWorldInData.org. Available at <https://ourworldindata.org/trust>.
- Richerson, Peter J., and Robert Boyd. 2005. *Not by Genes Alone*. Chicago, IL: University of Chicago Press.
- Rose, David C. 2011. *The Moral Foundation of Economic Behavior*. New York: Oxford University Press.

- Rose, David C. 2016. "Virtues as Social Capital." In Jennifer A. Baker and Mark D. White (eds.), *Economics and the Virtues: Building a New Moral Foundation* (New York: Oxford University Press), pp. 202–16.
- Rose, David C. 2019. *Why Culture Matters Most*. New York: Oxford University Press.
- Smith, Adam. 1759. *The Theory of Moral Sentiments*. D.D. Raphael and A.L. Macfie (eds.). Indianapolis: Liberty Fund Press (1982 edition).
- Smith, Adam. 1776. *An Inquiry into the Nature and Causes of the Wealth of Nations*. R.H. Campbell and A.S. Skinner (eds.). Indianapolis: Liberty Fund Press (1981 edition).
- Smith, Vernon L., and Bart J. Wilson, Bart. 2017. "Sentiments, Conduct, and Trust in the Laboratory." *Social Philosophy & Policy* 34: 25–55.
- Wicksteed, Philip H. 1933. *The Common Sense of Political Economy and Selected Papers and Reviews on Economic Theory*. London: Routledge & Kegan Paul.
- Williamson, Oliver. 1993. "Calculativeness, Trust, and Economic Organization." *Journal of Law & Economics* 36: 453–86.
- Wilson, Bart J. 2010. "Social Preferences Aren't Preferences." *Journal of Economic Behavior & Organization* 73: 77–82.
- Yamagishi, Toshio. 2000. *Trust*. Boulder, CO: Westview Press Incorporated.

C. ETHICS IN SCHOOLS OF ECONOMICS

CHAPTER 9

HUMAN ETHICALITY

Evidence and Insights from Behavioral Economics

SANJIT DHAMI AND ALI AL-NOWAIHI

9.1 INTRODUCTION

THE dominant paradigm in economics, neoclassical economics, is based on the *homo economicus* model. Fictional analogues of humans in this model, *Econs*, are assumed to be *amoral* and entirely *self-regarding*, devoid of any intrinsic morality, or any desire to behave in an ethical manner. Econs do not exhibit an *intrinsic preference* for honesty; truth-telling; keeping promises; trusting others and being trustworthy; reciprocating kind and unkind behavior of others; and caring about the fairness of procedures. Econs also have no feelings of remorse or guilt from letting down the expectations of others. Econs strive solely to maximize their own material well-being (*self-regarding* preferences) without regard to the well-being of others; that is, they lack *other-regarding* preferences.¹

Contemporary economics focusses, almost exclusively, on the behavior of Econs. The validity of this model is often taken as an article of faith among economists, which is likely to be surprising to other social scientists. However, when pressed, many economists are likely to express the view that they do not believe in the “literal truth” of such a model, but that such a model provides a “good approximation” to the real world. However, the “good approximation” part is never formally demonstrated.

Many economists are likely to argue that giving up a worldview based entirely on Econs will squander the hard-earned discipline of economic models and will open the way for a proliferation of ad hoc models. This is a common but deeply flawed argument. Discipline and progress in science arise from building models that are in conformity

¹ To be sure, neoclassical economics can be amended to include some forms of other-regarding preferences, such as “keeping up with the Joneses” or engaging in “snob or conspicuous consumption.” However, these features play, at best, a peripheral role in the actual practice of neoclassical economics, which is the benchmark that we are interested in.

with the empirical evidence and being prepared to amend models as new evidence emerges. As Richard Feynman, the Nobel Prize winner in physics, is once reported to have said: “We are trying to prove ourselves wrong as quickly as possible, because only in that way can we find progress.” If a model based on Econs is rejected, then alternative models that might be in better conformity with the evidence need not be ad hoc; for a more detailed discussion of these methodological issues, see the introductory chapter in Dhami (2016).

This chapter will argue that the *homo economicus* model is not consistent with the empirical evidence. Space limitations necessarily limit the discussion, but a longer and fuller treatment is available elsewhere (Dhami 2018). We will restrict ourselves to discussing issues of morality and ethics in this paper, with particular emphasis on the incentive to lie. Readers interested in the empirical validity of the self-regarding assumption behind the *homo economicus* model and the evidence for other-regarding preferences can consult several good sources, such as Camerer (2003), Fehr and Schmidt (2006), Dhami (2016: pt 2), and Gintis (2017).

There should be no presumption that intrinsic human morality and human virtues dilute the rationality assumption in economics. Rationality simply requires that people should have consistent preferences. The presence of other-regarding preferences still leads to rational choices in this sense (Dhami 2016: sect. 5.2.3). This chapter outlines the emerging evidence on ethics and morality in behavioral economics.

Neoclassical economics does not deny the existence of virtuous human behavior such as reciprocity, truth-telling, and keeping promises. However, it ascribes the cause of this behavior to *extrinsic preferences* for maximizing one’s own material well-being. Econs tell the truth when the current benefits from lying are lower than the discounted future costs of such lies. In strategic situations, Econs might be induced to reveal the truth because clever contracts guarantee lower payoffs from lying, relative to truth-telling; this approach lies at the heart of *contract theory* and *mechanism design*. The motivation to be intrinsically honest is not taken into account in these contracts because Econs are amoral. Rather, any observed virtuous behavior is assumed to be merely instrumental in increasing material well-being. Fortunately, a strength of the *homo economicus* model is that it typically makes precise, testable, predictions, so we can test whether morality is intrinsic or extrinsic and whether morality is influenced by factors such as the context, frame, the size of incentives, and competition. We shall argue below that the predictions arising from this framework are not well supported by the evidence.

Econs are *consequentialists*; they consider the material payoffs arising from their own actions but neglect the potentially harmful effects on others. The bulk of the empirical evidence indicates that the actions of players are *non-consequentialistic*. For instance, Gneezy (2005) finds that of the two following scenarios, selling a car with failing brakes or a car with broken pump that both cost the same to repair, the latter is considered more unethical.

Several decades of research in behavioral economics have shown that if the aim is to better explain and understand human behavior, then *homo behavioralis* is a superior

candidate to replace *homo economicus*. *Homo behavioralis* cares for material interests and for extrinsic incentives (just like *homo economicus*), but it also exhibits a strong sense of morality, considers the ethicality of alternative options, responds to intrinsic incentives, and is conditionally reciprocal. *Homo behavioralis* has been central to the development of behavioral and experimental economics, the fastest growing and most exciting development in economics in recent decades (Camerer 2003; Dhami 2016; Gintis 2017).

This chapter is divided into several sections that sometime overlap. Section 9.2 explains different canonical lab experiments that have been used to study lying behavior. Section 9.2.2 considers the field evidence, while Section 9.2.3 considers the external validity of lab evidence. Section 9.2.4 examines the effects of incentives on lying. Section 9.3 explores some of the microfoundations for moral behavior. Section 9.4 considers many different aspects of moral behavior, such as those arising from gain and loss frames, delegation, third party punishment, moral suasion, and social identity. Section 9.5 introduces how one might make use of psychological game theory to consider emotions such as guilt, shame, and intentions, that may underpin moral behavior. Section 9.6 examines the relation between markets and morality, and Section 9.7 concludes.

9.2 EVIDENCE ON LYING

Consider some useful terminology (Gneezy 2005; Erat and Gneezy 2012). Econs feel no remorse, disutility, or guilt, in telling *selfish black lies*—lies that benefit the liar but potentially harm others. *White lies* benefit others, but they must not decrease the liar's own utility. *Altruistic white lies* may harm the liar but benefit others. Finally, *Pareto white lies* improve the utilities of the liar and of others. There are two main experimental methods for studying lying behavior; we consider these next.

9.2.1 Experimental Methods

9.2.1.1 Lying Behavior in Sender-Receiver Games

In *sender-receiver games*, player 1 sends a message to player 2, who observes the message, and then takes an action that yields payoffs for both players. Consider the following three treatments in Gneezy (2005) in Table 9.1 in a sender-receiver game.

Player 1 (the sender) can observe all the payoffs in Table 9.1, but player 2 cannot. Player 1 can send two possible mutually inconsistent messages, A and B, to player 2. Message A (respectively, B) says: “Option A [respectively, B] will earn you more money.” Player 2 always earns more from option A, so message B is a lie but the sender always receives a higher payoff from option B. Player 2's chosen option is implemented.

Table 9.1 Payoffs of the players in each treatment in Gneezy (2005). The percentage of senders who choose each option is also shown in the brackets

Treatment	Option	Payoff of	
		Player 1	Player 2
1	A (64%)	5	6
	B (36%)	6	5
2	A (83%)	5	15
	B (17%)	6	5
3	A (48%)	5	15
	B (52%)	15	5

In an *incentive compatible elicitation*, 82 percent of the senders expected the receiver to trust their message; in actual practice, 78 percent of the receivers trusted the messages. The incentive to lie for the sender differs in the three treatments. First, the increase in the sender's payoff from telling a lie is 1 for treatments 1 and 2 but it is 10 for treatment 3. Second, the receiver's loss from being lied to is 1 in treatment 1 but 10 in treatments 2 and 3. Both these incentives play an important role in the sender's decision. The percentage of senders who lied in treatments 1, 2, and 3, respectively, is 36, 17, and 52, and these percentages are statistically different. However, a potential confound in the experiment is that it cannot cleanly separate the *role of other-regarding preferences* from *aversion to lying*.

9.2.1.2 Lying Behavior in the Die-in-the-Cup Paradigm

Fischbacher and Föllmi-Heusi (2013) introduced a novel method of measuring aggregate lying, the *die-in-the-cup* paradigm (or the *die-cup experiment*). Subjects roll a six-sided die in private, observe the outcome, and make a report of any number between 1 to 6. For the reported numbers 1, 2, 3, 4, and 5, the payoff equals the number, but for a report of the number 6, the payoff is zero.

Because outcomes are only privately observed, self-regarding and amoral Econs should report the number that achieves the highest payoff—that is, 5. *Individual lying* cannot be verified, but the distribution of reports can be compared with the predicted statistical distribution under truth-telling (each of the numbers 1–6 has a probability 1/6, or 16.7 percent) to determine the extent of *aggregate lying*. These experiments rule out altruism and reputational concerns.

Figure 9.1 is based on 389 participants. The stars on the top of each percentage sign on a histogram bar denote the 1% significance levels for a two-sided binomial test of the differences between the private reports and the statistical prediction under truth-telling.

A Kolmogorov-Smirnov test rejects equality of the distribution of private reports and the predicted distribution under truth-telling, thus, there is significant aggregate

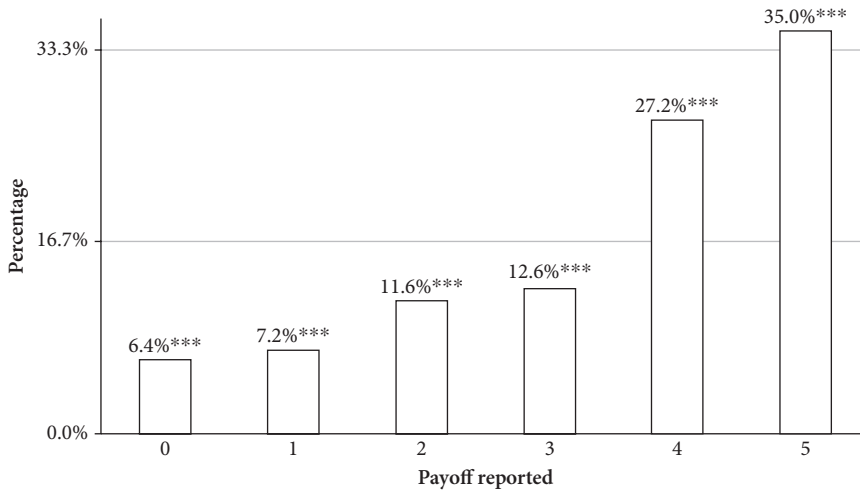


FIGURE 9.1 Results of the baseline treatment in the die-cup experiment.

lying. Numbers 1, 2, 3, and 6 that receive relatively low payoffs, are underreported, while numbers 4 and 5 are overreported. Statistically, only 16.7% of the subjects are predicted to get a 6 under truth-telling. Econs should never report a 6, yet 6.4% of the subjects report 6; because $\frac{6.4}{16.7} = 0.384$, the proportion of intrinsically honest subjects is 38.4%. Statistically, 16.7% of the subjects would have got a 5 anyway, and the percentage reporting 5 is 35%; hence, the percentage of people who are unethical income maximizers is $\frac{6}{5}(35 - 16.7) = 21.96\%$. Further, 27.2% of the subjects report the number 4, yet only 16.7% should get a 4 under truth-telling. Hence, some subjects do not lie maximally; these *partial liars* could be motivated by a desire to maintain a positive self-image, or they may suffer costs of lying.²

9.2.2 Evidence from the Field

In a field study in Germany, conducted by Abeler et al. (2014), participants anonymously tossed a coin in the privacy of their homes, and received 15 euros if they reported tails, and nothing otherwise. The percentage reporting tails is lower than 50 percent, suggesting almost complete truth-telling, despite the substantial inducement to lie. The relatively greater honesty in the home environment relative to the lab suggests that lying in a home or family setting may be considered more unethical.

² Abeler et al. (2016) conduct a meta-analysis of 90 studies that use the Fischbacher-Föllmi-Heusi framework. They found that, on average, subjects forgo three quarters of the potential gain from lying; this could either be due to honest reporting, or due to partial lying. This result is robust to increasing the payoff levels.

Pruckner and Sausgruber (2013) conducted field experiments in two towns in Austria, in which a newspaper, costing 0.60 euros, is sold on the street using a booth filled with newspapers. Customers pay into an unmonitored padlocked cashbox. There is very little material gain from not paying, and a very small probability of non-payments being discovered. When subjects are primed for a moral norm of payment through a posted note on the cashbox (as compared to a note stressing a purely legal obligation to pay), self-interest is reduced somewhat in favor of a social norm of honesty that requires payment for the newspaper.

Utikal and Fischbacher (2013) apply the die-cup experiments to a sample of Franciscan nuns in Germany and compare the results with a student subject pool. Students overreport in a manner consistent with the results of die-cup experiments. However, nuns tell *disadvantageous lies* that harm them but benefit others (altruistic white lies). No nuns report the numbers 4 and 5; perhaps they believed that (honestly) reporting these numbers might invite a suspicion of dishonesty.

Azar et al. (2013) observe whether clients at an Israeli restaurant return (deliberately given) excess change. They find that long-term clients at the restaurant are more likely to return the extra change, although a majority of the clients (128 out of 192) do not return the change. When the amount of extra change is increased, surprisingly, the authors find that it is more likely to be returned. The authors conjecture that a greater psychological cost must be paid to keep a larger amount of extra change. A potential confounding factor is that subjects might not notice small amounts of excess change, preventing the option of returning this change.

9.2.3 External Validity of Lab Evidence

There have been concerns about the external validity of lab experiments. But the emerging consensus is that lab evidence has a high degree of external validity that is possibly no different from the external validity of field experiments themselves (Camerer 2015; Dhami 2016).

Dai et al. (2018) find that those who have just been caught evading fare in public transport behave honestly in the lab experiment. The distribution of their reports is statistically indistinguishable from ticket holders; *conscience-accounting* may be an explanation. When the die task is replaced by a contextualized lab public transport game that allows for fare evasion, then self-reported fraudsters in the field also behave more dishonestly than the rest. Overall, lab behavior has a high degree of external validity.

Alm et al. (2015) compare the lab behavior of student and non-student populations in a tax evasion experiment in which subjects are informed about the probability of an audit and the penalty rate. Mean compliance levels differ. However, the distributions of the compliance rates and the behavioral responses to changes in the compliance parameters for the two groups are statistically identical. Armantier and Boly (2013) consider a corruption experiment (candidate offers a bribe to a grader to increase the grades) in

three different settings: a lab in a developing country, a lab in a developed country, and in the field in a developing country. They find similar qualitative and quantitative results when they compare the three different groups of subjects.

In Hanna and Wang (2017), 699 students in India play a modified die-cup experiment consecutively over 42 times. The resulting distribution, for each student, can then be compared against the statistical uniform distribution under truth-telling to reveal if each student lied or not. The students were then asked if they prefer government/public-sector jobs. Those who reported above-median scores on the die were also 6.2 percent more likely to choose such jobs that are often associated with greater corruption. A similar association is found between those who have less pro-social preferences, as revealed by their actions in a dictator game.

9.2.4 Incentives and Lying

In sender-receiver games, the incentive to lie influences the extent of lying, suggesting that incentives may loosen one's morality (Gneezy 2005; Erat and Gneezy 2012). In other methods, such as the matrix task and the die-cup experiments, cheating does not appear to respond to the extent of incentives.³ What accounts for this difference?

In a matrix solving task, Mazar et al. (2008) find that at low incentive levels (\$0.10, \$0.50 per correct matrix), there is a small level of dishonesty, but at higher incentive levels (\$2.50 and \$5.00 per solved matrix), there is no lying. This suggests a convex cost of lying, so that when lying for higher amounts, the marginal cost of lying to an individual increases. Using the die-cup experiment, Fischbacher and Föllmi-Heusi (2013) find that even when incentives are tripled, the extent of lying does not change. These general findings are supported in the meta-analysis of Abeler et al. (2016). However, when subjects are told that lying is legal, then the extent of lying responds to incentives (Gibson et al. 2013).

Kajackaitea and Gneezy (2017) try to reconcile these results by proposing the *mind game*. Subjects first privately think of a number between 1 and 6, then they privately roll a six-sided die. If they self-report a congruence between the two numbers, then they receive a prize with different incentives (\$1, \$5, \$20, \$50). They term the classical die-cup experiment, described above in Section 9.2, the *cheating game*. There is greater lying in the mind game relative to the cheating game for all levels of incentives. There is no trend in lying as incentives vary and the lowest level of lying occurs when the incentives are the highest (perhaps suggesting a convex cost of lying). The difference in cheating rates between the stake sizes \$5 and \$50 is statistically significant, although none of the other pairwise comparisons of lying behavior has statistical significance.

³ In a matrix task subjects are given a timed mathematical problem. For instance, detecting the number of pairs of numbers from a given set of numbers in a matrix that add up to 10. Subjects are then given the correct answers and have to self-report the number of their correct answers.

9.3 SOME EXPLANATIONS FOR MORAL BEHAVIOR

In this section, we consider several explanations for why people depart from the assumption of Econs in the *homo economicus* model.

9.3.1 Maintenance of Self-Image

Mazar et al. (2008) proposed a theory of self-concept, or self-image maintenance, which has well-established antecedents in psychology. In this theory, people have in mind some reference standard of behavior, say, relating to the desired degree of honesty. This reference standard, possibly context dependent, could conceivably be influenced by social norms for such behavior, or by one's own internal moral compass. When individuals take an action that falls below the reference standard, then they negatively update their self-image, which is aversive. Conversely, when actions exceed reference standard, individuals might positively update their self-image. These standards of behavior may be clearly categorizable on some ethical criteria, or they may fall into ambiguous categories, in which case, one may engage in self-serving justifications and rationalizations of having met the standards.

Self-image may be malleable to the extent that one can be dishonest up to a limit without having any adverse effect on one's self-image. But as this limit is exceeded, individuals negatively update their self-image. This could explain, for instance, why people engage in partial lying instead of maximal lying, a common finding in the die-cup experiments. Splitting the benefits of a dishonest action with others may also produce a less negative update to the self-concept, as does the telling of white lies (Erat and Gneezy 2012; Gino et al. 2013).

9.3.2 Cost of Lying

Individuals could incur a direct cost of lying, that is, a direct moral cost that is subtracted from the utility they derive from an action. Or it could be that they suffer psychological costs such as guilt-aversion that inhibit lying (Dhami 2016: 454), costs that arise from letting down the expectations of others that one is honest; that is, they hinge on the second order beliefs of players. Abeler et al. (2016) find evidence for a direct preference for being honest and for being seen to be honest. One may also be subject to the trade-off between getting higher material payoffs through lying and reduced utility on account of moral transgressions and a deterioration in self-image; this has been termed as *ethical dissonance* (Barkan et al. 2012).

Using a sender-receiver game in Bangladesh, Leibbrandt et al. (2017) introduce the option to remain silent in addition to sending either a true message or a false message. At high levels of stakes, worth about several months' average wage, they find that this reduces the likelihood of sending a true message by 30 percent. However, there is no difference in the likelihood of sending a false message, thus, the option to remain silent is often exercised. These data are not consistent with the theory that people are intrinsically honest, but rather that there is a cost of lying that is balanced against the benefit of being virtuous. Another interesting feature of this dataset, as compared to the datasets from sender-receiver games in the Western world, is that only 54 percent of the receivers actually follow the sender's recommendations (compare this with 82 percent in Gneezy 2005). Senders were significantly more optimistic and believed that 67 percent of the receivers would follow their recommendations; although 55 percent of the messages were true, receivers, on average, believed that 48 percent would be true, indicating that they probably took account of the option of remaining silent.

Gneezy et al. (2018) propose to disentangle the intrinsic cost of lying into several components. The direct cost of lying depends on two components: a fixed component that is independent of the size of the lie, and a variable component that depends on the magnitude of the lie. In addition to the direct cost, another cost arises from social identity concerns that capture how one's degree of honesty might be perceived by others. This too has a fixed component and a variable component that is determined by the (endogenously determined) probability that others believe one's actions to be honest. This framework can explain partial lying, and has empirically supported comparative statics with respect to changes in the extent of lying as (1) the objective probability of the underlying outcomes changes, and (2) one varies the degree of observability of the lie by others.

For an attempt to explain the data in Fischbacher and Föllmi-Heusi (2013) using both the concerns mentioned so far—that is, image concerns and a cost of lying—see the theoretical model in Khalmetski and Sliwka (2017).

9.3.3 Moral Balancing

Gneezy et al. (2014) provide another explanation for moral actions. Individuals might have self-imposed moral standards of behavior. If these are transgressed, then the individual might wish to engage in compensatory behavior, say, on account of guilt, to undo past transgressions (*moral balancing*). The authors give two nice examples of religious practices that tap into such a desire: the practice of *ashamot* (guilt) offerings as atonement for past transgressions in the Jewish faith, and *tariff penances* that date back to the medieval Catholic Church. Immediately following a transgression, in the hot state, one experiences a high level of guilt, but guilt depreciates over time as one enters a cold state. Individuals may also be forward-looking and take account of the subsequent depreciation of guilt, in the process adjusting the level of their current

transgressions. These ideas are developed and confirmed by Gneezy et al. (2014) by adapting the sender-receiver experiments of Gneezy (2005) described above; a dishonest message constitutes a moral transgression.

Ploner and Regner (2013) ask subjects to first privately roll a six-sided die. Subjects who report an odd (even) number are entitled to play a dictator game with a higher (lower) endowment. The statistical probability of each report under truth-telling is $\frac{1}{2}$; however, significantly more than 50 percent of the subjects report an odd number. Dictators who lied earlier (by falsely claiming an odd number) transfer more money to the receiver relative to a baseline dictator game with an identical endowment where there is no possibility to cheat (the die is rolled in the presence of the experimenter). This suggests moral balancing.

Confessions may also be thought of as a form of moral balancing, in which one confesses to reduce guilt from an immoral act, or simply to morally cleanse oneself. However, evidence suggests that like partial lying in the die-cup experiments, subjects engage in partial confessions, taking only *partial blame* for immoral acts (Peer et al. 2014).

9.3.4 Self-Serving Justifications

People might engage in self-serving justifications to weaken the apparent immorality of their actions, particularly when the moral benchmark might not be fully clear (Shalvi and Leiser 2013). For instance, when people privately roll a die in the die-cup experiments, they report lower numbers relative to a treatment in which they can roll the die thrice but are asked to report only the outcome for the first throw. The extra throws should be irrelevant for someone who wishes to tell the truth. However, for a potential liar, the situation is different. It might be considered more morally justifiable to report a higher number taken from the second or third throws of the die, even though one is asked to report the number only on the first throw. Similarly, when individuals can justify immoral actions that benefits others, they are more likely to lie (Conrads et al. 2013).

9.3.5 Public Personas, Private Personas, and Morality

Morality is likely to be context dependent. Gintis (2017: ch. 3) distinguishes between the *private persona* and the *public persona* of individuals that arise, respectively, in the private sphere and the public spheres of their actions. In the private sphere, individuals engage in private everyday transactions, such as “how much to save” and “when to retire.” The public sphere is, according to Gintis, “the locus of activities that create, maintain, transform, interpret, enforce, and execute the rules of the game that define society itself” (ibid.: 47), including voting in elections and participating in a civil rights movements. Actions in the public sphere are non-consequentialistic; they give rise to no private material payoffs, nor does any individual action, on its own, alter social outcomes.

Individuals appear to behave “as if” they possess different preferences in the private and public spheres. In the private sphere, and under self-regarding preferences, individuals have private personas that are predicted to behave like the Econs in neoclassical economics. However, in the public sphere, individuals appear to have public personas that derive direct utility from participating in actions in the public sphere. For instance, individuals might derive direct utility from voting in elections or from participating in social movements. However, such a preference is not absolute. Individuals are likely to weigh the extra utility from these actions against the extra cost.

Thus, much of rational choice theory that is devoted to making sense of voting and participation in social actions is simply based on an incorrect assumption: namely, that individuals take purely consequentialistic actions by engaging their private personas. Gintis suggests that the appropriate equilibrium notion in the public sphere is a form of *social rationality*, as encapsulated in a Kantian equilibrium. In a symmetric n -player game, a Kantian equilibrium strategy is such that every player prefers it to all other strategies if “everyone who shares their preferences were to act according to the same rule” (Gintis 2017: 51).

Dhami and al-Nowaihi (2010a, 2010b) consider a theoretical model of behavioral political economy in which voters have other-regarding preferences (as in Fehr and Schmidt 2006). They show that such voters behave in a self-regarding manner when choosing their individual labor supply, but behave in an other-regarding manner when choosing societal redistribution through voting. This is observationally equivalent to having a private persona in one sphere and a public persona in the other and as such provides some microfoundations to the idea. An analogy might help: a rich voter may send his own children to a private school, but also vote for more public funds for state education. Dufwenberg et al. (2011) showed that this feature applies to a more general class of social preferences (see Dhami 2016: sect. 6.5, 6.6).

9.4 EXPLORING THE RICHNESS OF HUMAN MORALITY

This section considers a wide variety of issues that arise in the study of human morality. These include heterogeneity of lying behavior; differences in lying in gains and losses; effects of history, groups, delegation, moral suasion, and social identity on lying; cross-country differences in lying; and gender differences in lying.

9.4.1 Types of Liars

It appears that there is a spectrum of lying types; Gneezy et al. (2013) identify eight such types. Some subjects are always honest, irrespective of the incentives. Others always

maximize their monetary benefits (similar to the amoral, and self-regarding preferences in neoclassical economics); the respective percentages of these two groups among the senders of the message in a sender-receiver game are 33 percent and 28 percent. Another group responds to the incentives to lie-lying more when the incentives to do so are high. Interestingly, as subjects gain experience of lying, they lie more.

9.4.2 Lying in Gain and Loss Frames

Morality might also be influenced by whether, relative to some *reference point*, one is in a *gain frame* or in a *loss frame*. This is because, due to loss aversion, losses typically bite, on average, about 2.25 times equivalent gains—a robust finding in humans and close primate relatives (Dhmi 2016: chs 3, 20, 21).

Schindler and Pfattheicher (2017) use a variant of the die-cup experiments. They ask subjects to privately roll a die 75 times and report the number of occurrences of getting a 4. Since the probability of a 4 in each throw is $1/6$ and the throws are independent, the statistical prediction under truth-telling is $75(1/6) = 12.5$. In the gain frame, subjects gain 10 cents for each reported 4. In the loss frame, subjects are initially endowed with 7.5 euros and told that they will lose 10 cents for every report of a number that is not 4. This is identical to the opportunity cost of not reporting a 4 in the gain frame. On average, subjects in the loss frame reported significantly higher 4s as compared to truth-telling ($p = 0.031$), which indicates significant dishonesty. In contrast, in the gain frame, no statistically significant dishonesty was found.

Grolleau et al. (2016) give subjects a matrix solving task in a 2×2 design (gain versus loss frame and monitored versus unmonitored reporting). In the gain frame, subjects are given a payment for the number of correct solutions. In the loss frame, subjects are initially given an endowment and payment is clawed back from them for the unsolved matrices. Relative to the monitored frame, in the unmonitored frame where misreporting of the number of correct matrices is possible, the percentage of solved matrices increases by 43 percent in the gains frame and by 296 percent in the loss frame. Both differences are significant, and significantly different from the monitored frame.

Several other papers also report an increase in cheating when subjects are in the loss frame relative to the gain frame. The opportunity to convert a loss into a gain induces greater cheating (Shalvi 2012). Goals might serve as a reference point and falling below goals presumably puts people in a loss frame (Dhmi 2016: pt. 1, sect. 3.7). It is found that there is more unethical behavior when subjects fall below their goals (Schweitzer et al. 2004).

Suppose that there are only two outcomes, $x_1 < x_2$, received with respective probabilities p and $1 - p$ ($p \in (0, 1)$). The decision-maker has the option to report either of the two outcomes in a truth-telling task. Assume that the reference point, in the sense of prospect theory, is the average outcome, $\bar{x} = px_1 + (1 - p)x_2$, where $x_1 < \bar{x} < x_2$. An increase in p increases the reference point $\bar{x}$; thus, reporting the lower outcome will reduce the

individual's utility proportionately more on account of loss aversion. Thus, the incentive to lie is sensitive to the probability distribution of outcomes. Garbarino et al. (2017) experimentally confirm this prediction.

9.4.3 Delegation and Punishment

In experiments conducted in the West, punishments are typically pro-social; that is, cooperators punish non-cooperators (Fehr and Gächter 2000). However, in the rest of the world, such punishment can take an anti-social form (non-contributors punish contributors as revenge for past punishments). Norms of civic cooperation (such as attitudes toward tax evasion and abuse of the welfare state) and the rule of law (such as people's trust in law enforcement institutions) are positively correlated with pro-social punishment (Herrmann et al. 2008). Humans also appear hardwired to inflict third party punishment on others; this is an essential component in the maintenance of human morality and social norms (Gintis 2009, 2017; Dhami 2016: pt. 2).

In Bartling et al. (2014), dictators have either a beneficial or a harmful effect on receivers (say, in terms of the amount given to the receiver). Dictators can, costlessly and voluntarily, choose to be informed, or remain ignorant of the effect on the receivers. A third party observes the choice of the dictator and decides whether or not to undertake costly punishment of the dictator. Third party Econs would never engage in such punishment because bygones are bygones (Dhami 2016: sect. 6.2, p. 52). Conditional on the dictator having chosen to remain ignorant and a harmful effect on the receiver, the dictator is punished less. Thus, ignorance helps to reduce the blame for an unfair outcome. However, third party punishments on the dictators when the outcome is beneficial and the dictators choose to be ignorant (relative to being informed) are significantly higher. Thus, on net, it might not help dictators to stay ignorant.

Bartling and Fischbacher (2012) show that blame for unpleasant outcomes can be successfully shifted by delegating decisions in a dictator game to a third party. Conditional on an unfair outcome, the average punishment meted out to the dictator is 4.27 points in the absence of delegation and 1.13 points in the presence of delegation—a statistically significant difference. Unlike Blount (1995), even delegation to a random device appears to shift some of the blame for an unfair outcome away from the dictator. When the dictator cannot choose an unfair outcome but an unfair outcome can only result from delegation, the dictator is punished relatively more. Perhaps this allows third parties to better infer the unkind intentions of the dictator. In a nutshell, the moral aspects of behavior help explain the pattern of delegation.

Erat (2013) finds that people prefer to delegate lying to others. Further, people delegate more as the harm to others increases; in this case, women are even more likely to delegate than men. Whistle-blowers in organizations, a form of third-party punishment, often put themselves at substantial risk and reduce their own material rewards to report unethical behavior (Reuben and Stephenson 2013); this form of behavior is inconsistent with Econs.

9.4.4 History-Dependent Lying

One may cheat more if one feels unfairly treated in the past. In order to test this intuition, Houser et al. (2016) first play a dictator game with subjects who are randomly assigned to be either dictators or receivers. Having played this game, subjects are then given an ethical choice. They are asked to privately toss a coin; a report of heads earns them 1 euro and a report of a tails earns 3 euros. There is significant lying; 74.5 percent report tails, while under truth-telling 50 percent should report tails. Using a mixture model with two types, those who lie and those who do not, the authors estimate that the cheating rate among receivers is higher than that among dictators (53 percent versus 45 percent). Furthermore, this difference is driven almost entirely by those receivers who received nothing in the dictator game. There is no noticeable difference in cheating between dictators and receivers who received a positive amount, despite the dictator's payoffs being greater.

9.4.5 Lying in Groups

Groups may induce more or less honesty depending on several factors. (1) Groups may induce a more sophisticated analysis relative to individuals. (2) Individuals might be able to disguise their lying in groups. (3) Individuals might lie more in groups, on account of social preferences (such as lying for team members). (4) Concerns for one's social image might influence lying in groups, depending on whether the group is honest or not. (5) Group interaction may reveal social norms about honesty and dishonesty.

Kocher and Schudy (in press) use a variant of the die-cup experiments to study some of these factors. In the *individual treatment*, subjects observe the throw of a die on a computer screen and self-report the observed number for a reward. The experimenter also observes the outcome of the die on the computer screen, so individual cheating can be identified (which might deter cheating because subjects know this). Subjects then participate in either of two group cheating tasks where they have the opportunity to chat and exchange free-form text messages before they individually submit their reported numbers.

There is significantly more lying in the two group tasks (89.7 percent and 86.3 percent) compared to the individual treatment that precedes it (61.5 percent). What causes increased lying in groups? First, group chat increases the beliefs of players about the dishonesty of others, relative to the individual treatment. Second, the nature of communication during the group chat plays a key role: 43.4 percent of the arguments favor dishonesty while 15.6 percent favor honesty. Arguments for honesty significantly reduce lying in groups. These results suggest that there is a shift in the perception of individuals in groups about the honesty norms in the rest of the population that drives the results.

Balafoutas et al. (2017) conducted a real effort experiment with professional German internal auditors who received points for identifying the number of incorrect calculations

in a given task. Under *individual incentives*, subjects received a piece rate of 2 euros for each point. Under *competitive incentives*, random groups of two individuals were formed and the individual who gives more correct answers gets 4 euros for each point, whereas the partner gets nothing. Under *team incentives*, each team player gets 1 euro for a correct answer given by any of the team members. The number of correct answers was determined either by the individual, by the peers, or objectively by the experimenter, depending on the treatment.

Under individual incentives and peer evaluation, the actual and reported performance are very similar; thus, unless peers have a monetary incentive to misreport, they report relatively accurately. Under objective evaluation and team incentives, players underreport the performance of the team member, despite their report having no bearing on the payoffs, perhaps to enhance their self-image and status perception.⁴ Under peer evaluation and competitive incentives, there is underreporting of the number of correct answers achieved by the partner. The opposite—overreporting—occurs under team incentives. Dishonest behavior is driven by a minority of the subjects, while most subjects are honest. The share of truthful reports ranges from 70 percent (under peer reporting and team incentives) to 86 percent (under peer reporting and individual incentives).

9.4.6 Moral Suasion and Morality

People may be induced to act morally through simple *moral suasion*. Indeed, an older literature in the economics of banking stressed the positive role of moral suasion in banking, whereby the Federal Reserve in the United States used “persuasion” (such as closed door meetings with bank directors and appealing to the “public-spiritedness” of actions) to informally regulate private banks as compared to formal and binding regulation. Moral suasion is not predicted to have any effect on Econs unless it reveals some relevant information or if there are reputational issues at stake.

In two natural field experiments, Hallsworth et al. (2017) randomized five different messages across 100,000 taxpayers. A control group received a standard letter. In other treatments, letters were sent out where moral persuasion was involved; three of the messages were norm-based, while the remaining two were public service messages (such as “taxes fund public services”). Relative to the control treatment, the five messages resulted in an increase in the likelihood of an earlier repayment of taxes. The most successful of these messages produced a treatment effect of 5.1 percent over the control treatment.

⁴ Carpenter et al. (2010) and Charness et al. (2014) also find evidence that in the absence of monetary incentives, people who compete with each other may sabotage each other. Faravelli et al. (2015) also find that under competitive incentives, self-reported performance increases, as does the proportion of individuals who self-select themselves for such tasks.

9.4.7 Morality and Social Identity

There has been an explosion of research on social identity in economics (Dhami 2016, ch. 7). People identify themselves with social categories, groups, or identities, each having its own norms and ideal behaviors. One observes favorable behavior towards ingroup members (members of the group) and unfavorable behavior towards outgroup members (non-members of the group). What may constitute immoral and unacceptable behavior towards ingroup members might be perfectly acceptable when directed towards outgroup members. Killing of an ingroup member would, in most groups, be considered immoral, yet in times of war, perhaps aided by propaganda, killing of outgroup members is considered valiant and praiseworthy.

Humans appear so hardwired to respond to ingroup/outgroup distinctions in behavior that when they are primed for even minimal group identities, such as red and blue groups, they favor ingroup members relative to outgroup members. The presence and persistence of stereotypes and of discrimination towards other groups can be explained along these lines. Mullainathan and Bertrand (2004) respond to “help wanted” ads in Boston and Chicago newspapers, and for identical resumes, randomized among African American and white sounding names, the latter receive 50 percent more callbacks. In the trust game experiments of Eckel and Petrie (2011), subjects can view a photograph of the other player, at a price. Among those who chose to view the photograph, white trustors discriminate favorably towards white relative to black trustees. Black trustors do not discriminate, perhaps because many of their role models, such as schoolteachers and doctors, are white. On the whole, there appears to be an information value in a face that may be explained in terms of social identity.

Human morality is also tied to one’s *social* or *professional identity*. Cohn et al. (2014) divide 200 bank employees into a control group and an experimental group. Subjects privately toss a coin and, based on their reports, they can increase their income by up to \$200. The control group was primed for its non-occupational identity (such as “What is your favorite activity in your leisure time?”), and the experimental group was primed for its professional or occupational identity (such as “What bank do you work at?” or “How long have you been working in the banking sector?”).

In the control group, bank employees were honest. However, subjects in the experimental group are significantly more dishonest compared to those in the control group; 58.2 percent reported a successful coin flip, which is significantly different from the predicted 50 percent under truth-telling, and from 51.6 percent in the control group. For the non-banking employees, there was no difference in the honesty levels between the control and experimental conditions. The results are likely to be driven by the existing banking culture which appears not to be fully honest (subject to the usual caveat that the employees in experiment were not a random sample of all possible bank employees).

Conducting experiments on prison inmates in a Swiss prison, Cohn et al. (2013) show that when the subjects are primed for criminal identity, 32 percent lied, which is significantly different from the control condition. Further, the authors find a positive correlation between lying and lack of compliance of prisoner’s with prison regulations (such as aggressive behavior towards others and use of illegal drugs).

9.4.8 Morality and Anonymity

Immoral actions may be underpinned by emotions such as shame that arise from our actions being observed by others. In dictator games, Haley and Fessler (2005) show that when dictators are shown pictures of eyes in the same room while making decisions, they make more generous decisions. Perhaps the pictures induce a feeling of being watched, hence triggering emotions such as shame, if dictators make low offers.

Some people may cross over to the other side of a road when they see a beggar to avoid the guilt that they would feel if they did not give something to the beggar.⁵ Many dictators with an endowment of \$10 are found to exercise an option to exit the experiment with \$9 and have their decision kept anonymous rather than earn \$10 by playing the experiment (Dhami 2016: sect. 5.2.2). This is sometimes termed as *moral wiggle room*. Thus, many people might be termed *reluctant altruists*. However, it is now well known that many of the results from dictator game experiments do not survive in the presence of strategic interaction (Fehr and Schmidt 2006; Dhami 2016: sect. 5.2.2; Dhami et al., forthcoming).

Gneezy et al. (2017) found that people are unwilling to share their (possibly negative) views on the attractiveness of other people, even if shading their views incurs a personal material cost. When asked to share their views under anonymity, subjects are relatively truthful. The authors conclude that people do not wish to be messengers of bad news.

9.4.9 The Slippery Slope of Dishonesty

Garrett et al. (2016) examine changes in dishonesty when the opportunity to be dishonest is repeated over time. In their setup, an advisor is asked to provide advice to an uninformed estimator about how much money there is in a jar. The advisor's payoff is increasing in the amount of money that he or she states in this advice. When this game is repeated, and only the advisor benefits from the dishonesty (self-serving dishonesty), the advisor is found to engage in ever-increasing levels of dishonesty. The authors term this phenomenon the *slippery slope of dishonesty*. However, when dishonesty is other-serving—that is, it benefits the estimator but not the advisor—this phenomenon is not observed. Hence, it is selfishness, not altruism, that gives rise to the slippery slope of dishonesty.

9.4.10 Cross-Country Differences in Honesty

Hugh-Jones (2015) performed a coin-flip experiment in fifteen countries. Richer countries were more honest, on average, as were countries with a greater percentage of Protestant

⁵ On the other hand, an important explanation for why people give to charities is that they derive a warm glow from the act of giving. There has been support for this idea in some experiments and in neuroeconomic studies. However, there is debate about the relative importance of “pure altruism” and “warm glow” in charitable giving.

subjects. However, the main correlate of honesty is pre-1950 GDP differences, but not differential growth in GDP since 1950.

Gächter and Schulz (2016) construct a prevalence-of-rule-violations (PRV) index for 159 countries. The PRV index captures three kinds of rule violations: political fraud (using an index of political rights), tax evasion (proxied by the size of the shadow economy), and corruption (derived from the World Bank's control of corruption index). Subjects rolled a six-sided die twice in the die-in-a-cup method and were asked to report only the privately observed outcome of the first throw. The distribution of reports are close to neither the full-honesty benchmark nor the full-dishonesty benchmark. However, for 13 out of 23 subject pools, behavior is consistent with a justified ethicality benchmark, in which it is more ethical to report outcomes from the second throw rather than lie about the outcome based on a single throw. The fraction of individuals who are always honest varies from 4.3 percent to 87 percent across countries. No gender effects across countries are found.

9.4.11 Gender Differences in Lying

Several studies find that women are more honest than men. The evidence comes from dishonesty tasks, tasks involving moral costs of lying, lying in sender-receiver games, self-reports of coin tosses, return of excess change in restaurants, and fare dodging in a field experiment; see Houser et al. (2016) for the references.

Houser et al. (2016) explore some of the microfoundations of gender differences in lying behavior. Parents (88 percent mothers) could lie about the outcome of two coin tosses in the presence and in the absence of their children, and rewards were either for the parent or the child. Under truth-telling, the claim rates were designed to be 25 percent. Parents lied less in the presence of their child. The greatest claim rates, 58 percent, occur when parents report privately but the reward is for the child. This is statistically higher than the average claim rate of 33 percent, averaged across all treatments. The claimed win rate in the presence of a daughter is 28 percent, close to the predicted rate of 25 percent under complete honesty. This is significantly lower than the claimed average win rate of 42 percent in the presence of sons ($p < 0.01$). It suggests that a potential explanation for lower dishonesty among women may lie in the manner in which they are socialized when young, relative to men.

9.5 MORALITY AND BELIEFS: PSYCHOLOGICAL GAME THEORY

Classical game theory is not well suited to considering the role of a range of emotions—such as guilt aversion, surprise seeking, reciprocity, malice, anger, and shame aversion—that underpin human morality and play a critical role in the development and upkeep of

social and moral norms. In recent years, rapid progress has been made in *psychological game theory* (PGT), which allows beliefs to directly enter into the utility function of players. For a treatment of psychological game theory and the references that we suppress here, see Dhami (2016: sect. 13.5).

The following example illustrates how the feelings of surprise and guilt may directly impart disutility.

Example 1: John frequently visits cities A and B, and he typically uses a taxi to get around. In city A, tipping a taxi driver is considered insulting, while in city B it is the norm to tip a publicly known percentage of the fare. Suppose that it is common knowledge that if taxi drivers do not receive a tip, they quietly drive away. In city A, John gives no tip, and feels no remorse from not giving it. However, in city B, the taxi driver expects John to give him a tip (taxi driver's first-order belief) and John believes that the taxi driver expects a tip from him (John's second-order belief). Based on his second-order belief, John cannot bear the guilt of letting the taxi driver down by not paying the tip. Thus, he tips every time he takes a taxi in city B. Clearly, John's utility appears to be directly influenced by his second order beliefs, which is an impossibility under classical game theory.

In Example 1, in city B, if John believes that the taxi driver has been particularly courteous and helpful, then he might tip him extra, on account of reciprocity, which is particularly convenient to model in PGT.

Battigalli and Dufwenberg (2007) proposed a formal approach to modeling guilt. They distinguish between two different emotions associated with guilt.

- (1) *Simple guilt* arises from falling short of the perceived expectations of other players. For instance, if in city B in Example 1, John believes that the taxi driver expects a 15 percent tip, yet pays only a 10 percent tip, then he may suffer from simple guilt, which directly reduces his utility.
- (2) *Guilt from blame* arises when one cares for the attribution of intentions behind psychological feelings such as guilt aversion or surprise seeking. John feels further guilt if he believes that the taxi driver believes that John under-tipped him knowingly rather than by accident. Inferring such intentions requires third and fourth order beliefs. (For an application, see Dhami et al., forthcoming.)

Typically, models in PGT restrict themselves to analyzing beliefs up to the fourth order because it does not seem compelling that most people have the cognitive ability to form beliefs of higher orders. In contrast, classical game theory makes the empirically rejected assumption that players can form beliefs up to any order; that there is common knowledge in the form of an infinite regress of beliefs; and that beliefs and actions are consistent with each other (Dhami 2016: pt. 4). Indeed, there is no concrete, empirically founded, explanation in classical game theory as to how such a state might be achieved.

The surprise-seeking motive arises from exceeding the expectations of others as perceived by a player through his or her second-order beliefs. For instance, in Example 1, in

city B, John may believe that the taxi driver expects a tip that is 10 percent of the fare, yet he may derive extra utility by offering instead a 15 percent tip (surprise-seeking motive) that puts a smile on the taxi driver's face. One may extend these beliefs to higher orders by factoring in the intentionality of the surprise-seeking motive.

Dhami et al. (forthcoming) extend the theoretical framework of psychological games to a two-player public goods game that takes account of guilt aversion, surprise-seeking, attribution of intentions, and reciprocity. In an induced beliefs design, they find that all the emotions that underpin human morality play an important role in explaining contributions in public goods games. In particular, guilt aversion is, by far, the predominant finding at the level of the individuals and for the aggregate data. They find that for at least 30 percent of the subjects, the attribution of intentions behind guilt aversion and surprise-seeking is statistically significant, although they cannot rule out this motive for the remaining subjects.

Dufwenberg and Dufwenberg (2018) propose an explanation for the data in Fischbacher and Föllmi-Heusi (2013) by postulating that potential liars feel disutility from lying if they perceive that the audience believes that they are lying. They term this tendency *perceived cheating aversion*. Future experiments can test explicitly for this explanation by increasing the degree to which one's lying can be discovered by the audience.

9.6 MARKETS, INCENTIVES, AND MORALITY

The relation between markets and human sociality is of topical interest. Consider first a few examples; for references, see Dhami (2016: ch. 8) and Bowles and Polania-Reys (2010). Framing interactions between subjects in terms of market terminology, such as sellers, buyers, and bargaining, can diminish moral considerations. When workers in a field experiment are given an in-kind gift (a water thermos) relative to an equivalent monetary gift of \$20, their effort level increases by 25 percent, even when they have no preference for the thermos over \$20. Implicit incentives, such as bonuses, may highlight the moral aspects of one's actions, while explicit incentives, such as performance based pay, may turn off the moral frame, or trigger *moral disengagement*. When dictators in dictator game experiments are made to earn their endowments, or similar entitlements are created, they offer lower amounts to the receiver.

Economists stress the role of *extrinsic motivation* that responds to purely extrinsic material incentives. Yet, increasingly, behavioral economics has highlighted the role of *intrinsic motivation* (clearly, in addition to extrinsic motivation) and provided persuasive theoretical frameworks to study its effect on economic behavior (Dhami 2016: sect. 8.5, 8.6). Individuals signal to themselves and others through costly actions, such as charitable giving, that they are good and moral people. Indeed, extrinsic incentives, by removing or reducing the opportunity to engage in such signaling, may even crowd out intrinsic motivation (Ariely et al. 2009).

Titmus (1971) found that individuals are more likely to donate blood when they do so voluntarily through intrinsic motivation, rather than in the presence of monetary incentives. Some of the more recent findings are as follows (see Dhimi 2016: sect. 8.6 for references). Swedish data showed strong gender effects of incentives for blood donation; a crowding-out effect for women but not for men. In contrast, data from the American Red Cross blood drives (single events that solicit blood donation) revealed that incentives crowd-in blood donations. However, a significant increase in donations from incentives, in this study, came from substitution effects arising from other spatially and temporally separated blood drives. No effects of incentives on blood donations for long-term committed donors was found.

Gneezy and Rustichini (2000) found that when fines were levied for late arriving parents in private day care centers in Haifa, they arrived even later—exactly the opposite conclusion that might be drawn from Econs. This occurs because the fine places an extrinsic value on late arrivals, substituting for the intrinsic motivation of parents to not delay the caregivers. Parents continued to arrive late, even when the fine was removed, suggesting long-lasting, and negative, effects on intrinsic motivation.

Falk and Szech (2013) give individuals a choice between taking 10 euros or saving the life of a young healthy mouse who might be expected to live for two more years. In the non-market condition, subjects made this choice as isolated decision makers. In the market condition, subjects could bargain with another subject (bilateral setting) or several other subjects (multilateral setting). If subjects successfully bargain, they get 10 euros, but the life of the mouse is lost. Otherwise (unsuccessful bargaining), they lose 10 euros, but the mouse is saved. The percentage of subjects who are willing to accept 10 euros rather than save the life of the mouse in the non-market, bilateral, and the multilateral treatments is 45.9, 72.2, and 75.9 (respectively). Thus, markets reduce morality.

This could be on account of three possible reasons. First, relative to the non-market treatment, the responsibility for killing the mouse is spread over greater number of people in the market treatments, reducing guilt from killing the mouse. Second, in the process of bargaining, one may observe that others are willing to trade or bargain, hence condemning the mouse; this might loosen one's own morality. Third, markets might draw attention to a non-moral frame by focusing on bargaining, negotiations, and competition.

A range of interesting economic issues involve negative externalities in which actions by one party cause harm or disutility to another party. The typical policy response in economics, corrective taxes, ensures that the involved parties internalize the private and social costs of their actions. However, if economic agents care directly for social responsibility, might they internalize negative externalities anyway? For instance, many corporations stress the idea of corporate social responsibility, which requires corporations to take account of the larger social interest, even at a cost in terms of private profits. Consumers too are often willing to pay extra for socially responsible products that do not involve child labor or cruelty to animals, or that are made with greener, more expensive, technologies.

Bartling et al. (2015) design experiments to consider these issues in market and non-market settings. Firms and consumers can exchange a costly externality-free product or a cheaper negative-externality-causing product. Consistent with the presence of social responsibility, the baseline condition quickly stabilizes at 45 percent of the products being externality-free. As expected, such products are sold at a higher price relative to the externality-causing product, but the price difference is lower than the extra production cost of these products. Thus, in equilibrium, both sellers and buyers share in the costs of being socially responsible in competitive markets. When more competition between the sellers is introduced, the price drops further, but social responsibility does not. However, when the cost of production of the externality-free product is raised (from 20 percent of the surplus to 80 percent of the surplus), the degree of social responsibility falls. Furthermore, the frequency of choices in the non-market setting that mitigate the negative effects on third parties (the analogue of negative externalities in the market setting) is relatively higher. Hence, markets do appear to reduce ethicality, which is consistent with the findings of Falk and Szech (2013).

Evidence collected from cross-cultural variation in the outcomes of the ultimatum game suggests that the two main factors that enhanced human sociality were market integration in the community (the predominance of buying, selling, and working for a wage) and the degree of cooperation in production (whether production is carried out on an individual basis or in a team). Indeed, these two factors alone explained 66 percent of the variation in outcomes in the ultimatum game.

Thus, *markets enhance sociality, even if they might diminish ethicality*. By not making this important distinction, researchers risk drawing erroneous conclusions. Ethicality does appear to be influenced by norms for ethicality, but this begs the interesting question of why there are norms for some types of ethical behavior but not others.

9.7 CONCLUSIONS

The *homo economicus* model is not supported by the empirical evidence. A rich body of evidence now provides strong support for a *homo behavioralis* model. Exclusive reliance on the *homo economicus* model is neither justified on empirical grounds, nor on the mistaken belief that it leads to a more parsimonious account of real phenomena. Indeed, in order to ensure consistency with the data, several auxiliary conditions must be invoked in neoclassical models to account for the missing motivations found in *homo behavioralis*—a sort of “missing variables bias” in the theoretical models. This does not necessarily make the *homo economicus* model simpler, more parsimonious, or more appealing. The lack of compliance with the empirical evidence is sufficient grounds for moving beyond the *homo economicus* model. The main aim of this chapter is to make a strong case for such a move. We now have better data on the richness of human morality. The theoretical models in behavioral economics have made commendable progress. This is likely to accelerate in the future.

ACKNOWLEDGMENTS

We are grateful for comments to Herbert Gintis, Bjorn Bartling, Gary Charness, Marie Clarie Villeval, and Martin Dufwenberg.

REFERENCES

- Abeler, Johannes, Anke Becker, and Armin Falk. 2014. "Representative Evidence on Lying Costs." *Journal of Public Economics* 113: 96–104.
- Abeler, Johannes, Daniel Nosenzo, and Collin Raymond. 2016. "Preferences for Truth-Telling." CEDEX Discussion Paper No. 2016–13.
- al-Nowaihi, Ali, and Sanjit Dhami. 2015. "Evidential Equilibria: Heuristics and Biases in Static Games of Complete Information." *Games* 6: 637–76.
- Alm, James, K.M. Bloomquist, and Michael McKee. 2015. "On the External Validity of Laboratory Tax Compliance Experiments." *Economic Inquiry* 53: 1170–86.
- Ariely, Dan, Anat Bracha, and Stephan Meier. 2009. "Doing Good or Going Well? Image Motivation and Monetary Incentives in Behaving Prosocially." *American Economic Review* 99: 544–55.
- Armantier, Olivier, and Amandou Boly. 2013. "Comparing Corruption in the Laboratory and in the Field in Burkina Faso and in Canada." *Economic Journal* 123: 1168–87.
- Azar, Ofer H., Shira Yosef, and Michael Bar-Eli. 2013. "Do Customers Return Excessive Change in a Restaurant? A Field Experiment on Dishonesty." *Journal of Economic Behavior and Organization* 93: 219–26.
- Balafoutas, Loukas, Simon Czermak, Marc Eulerich, and Helena Fornwagner. 2017. "Incentives for Dishonesty: An Experimental Study with Internal Auditors." Working Papers in Economics and Statistics, University of Innsbruck. No. 2017–06.
- Barkan, Rachel, Shahar Ayal, Francesca Gino, and Dan Ariely. 2012. "The Pot Calling the Kettle Black: Distancing Response to Ethical Dissonance." *Journal of Experimental Psychology: General* 141: 757–73.
- Bartling, Bjorn, Florian Engl, and Roberto A. Weber. 2014. "Does Willful Ignorance Deflect Punishment? An Experimental Study." *European Economic Review* 70: 512–24.
- Bartling, Bjorn, and Urs Fischbacher. 2012. "Shifting the Blame: On Delegation and Responsibility." *Review of Economic Studies* 79: 67–87.
- Bartling, Bjorn, and Yagiz Ozdemir. 2017. "The Limits to Moral Erosion in Markets: Social Norms and the Replacement Excuse." CESifo Working papers, Vol. 17, No. 93.
- Bartling, Bjorn, Roberto Weber, and Lan Yao. 2015. "Do Markets Erode Social Responsibility?" *Quarterly Journal of Economics* 130: 219–66.
- Battigalli, Pierpaolo, and Martin Dufwenberg. 2007. "Guilt in Games." *American Economic Review Papers and Proceedings* 97: 170–6.
- Benabou, Roland. 2013. "Groupthink: Collective Delusions in Organizations and Markets." *Review of Economic Studies* 80: 429–62.
- Blount, Sally. 1995. "When Social Outcomes Aren't Fair: The Effect of Causal Attribution on Preferences." *Organizational Behavior and Human Decision Processes* 63: 131–44.
- Bowles, Samuel, and Sandra Polania-Reyes. 2012. "Economic Incentives and Social Preferences: Substitutes or Complements?" *Journal of Economic Literature* 50: 368–425.

- Camerer, Colin F. 2003. *Behavioral Game Theory: Experiments in Strategic Interaction*. Princeton, NJ: Princeton University Press.
- Camerer, Colin F. 2015. "The Promise and Success of Lab-Field Generalizability in Experimental Economics: A Critical Reply to Levitt and List." In Guillaume R. Fréchet and Andrew Schotter (eds.), *Handbook of Experimental Economic Methodology* (Oxford: Oxford University Press), pp. 249–95.
- Carpenter, Jeffrey, Peter. H. Matthews, and John Schirm. 2010. "Tournaments and Office Politics: Evidence from a Real Effort Experiment." *American Economic Review* 100: 504–17.
- Charness, Gary, David Masclet, and Marie C. Villeval. 2014. "The Dark Side of Competition for Status." *Management Science* 60: 38–55.
- Cohn, Alain, Ernst Fehr, and Michel A. Maréchal. 2014. "Business Culture and Dishonesty in the Banking Industry." *Nature* 516: 86–9.
- Cohn, Alain, Michel A. Maréchal, and Thomas Noll. 2013. "Bad Boys: How Criminal Identity Salience Affects Rule Violation." *Review of Economic Studies* 82: 1289–308.
- Conrads, Julian, Bernd Irlenbusch, Rainer M. Rilke, and Gary Walkowitz. 2013. "Lying and Team Incentives." *Journal of Economic Psychology* 34: 1–7.
- Dai, Zhixin, Fabio Galeotti, and Marie C. Villeval. 2018. "Cheating in the Lab Predicts Fraud in the Field: An Experiment in Public Transportation." *Management Science* 64: 1081–100.
- Dhami, Sanjit. 2016. *The Foundations of Behavioral Economic Analysis*. Oxford: Oxford University Press.
- Dhami, Sanjit. 2018. "Human Ethics and Virtues: Rethinking the *Homo Economicus* Model." CESifo Working Paper.
- Dhami, Sanjit, and Ali al-Nowaihi. 2010a. "Existence of a Condorcet Winner When Voters Have Other-Regarding Preferences." *Journal of Public Economic Theory* 12: 897–922.
- Dhami, Sanjit, and Ali al-Nowaihi. 2010b. "Redistributive Policy with Heterogeneous Social Preferences of Voters." *European Economic Review* 54: 743–59.
- Dhami, Sanjit, Mengxing Wei, and Ali al-Nowaihi. In press. "Public Goods Games And Psychological Utility: Theory and Evidence." *Journal of Economic Behavior and Organization*.
- Dufwenberg, Martin, and Martin A. Dufwenberg. 2018. "Lies in Disguise: A Theoretical Analysis of Cheating." *Journal of Economic Theory* 175: 248–64.
- Dufwenberg, Martin, Paul Heidhues, Georg Kirchsteiger, Frank Riedel, and Joel Sobel. 2011. "Other-Regarding Preferences in General Equilibrium." *Review of Economic Studies* 78(2): 613–39.
- Eckel, Catherine C., and Ragan Petrie. 2011. "Face Value." *American Economic Review* 101: 1497–513.
- Engelmann, Jan B., and Ernst Fehr. 2016. "The Slippery Slope of Dishonesty." *Nature Neuroscience* 19: 1543–4.
- Erat, Sanjiv. 2013. "Avoiding Lying: The Case of Delegated Deception." *Journal of Economic Behavior & Organization* 93: 273–8.
- Erat, Sanjiv, and Uri Gneezy. 2012. "White Lies." *Management Science* 58: 723–33.
- Falk, Armin, and Nora Szech. 2013. "Morals and Markets." *Science* 340: 707–11.
- Faravelli, Marco, Lana Friesen, and Lata Gangadharan. 2015. "Selection, Tournaments, and Dishonesty." *Journal of Economic Behavior & Organization* 110: 160–75.
- Fehr, Ernst, and Simon Gächter. 2000. "Cooperation and Punishment in Public Goods Experiments." *American Economic Review* 90: 980–94.

- Fehr, Ernst, and Klaus Schmidt. 2006. "The Economics of Fairness, Reciprocity and Altruism: Experimental Evidence and New Theories." In Serge-Christophe Kolm and Jean Mercier Ythier (eds.), *Handbook of the Economics of Giving, Altruism and Reciprocity: Foundations, Volume 1* (Amsterdam: North-Holland), pp. 615–91.
- Fehr, Ernst, and Frederic Schneider. 2010. "Eyes Are on Us, But Nobody Cares: Are Eye Cues Relevant for Strong Reciprocity?" *Proceedings of the Royal Society B: Biological Sciences* 277: 1315–23.
- Fischbacher, Urs, and Franziska Föllmi-Heusi. 2013. "Lies in Disguise: An Experimental Study on Cheating." *Journal of the European Economic Association* 11: 525–47.
- Gächter, Simon, and Jonathan F. Schulz. 2016. "Intrinsic Honesty and the Prevalence of Rule Violations Across Societies." *Nature* 531: 496–9.
- Garbarino, Ellen, Robert Slonim, and Marie C. Villeval. 2017. "Loss Aversion and Lying Behavior: Theory, Estimation and Empirical Evidence." Mimeo.
- Garrett, Neil, Stephanie C. Lazzaro, Dan Ariely, and Tali Sharot. 2016. "The Brain Adapts to Dishonesty." *Nature Neuroscience* 19: 1727–32.
- Gibson, Rajna, Carmen Tanner, and Alexander F. Wagner. 2013. "Preferences for Truthfulness: Heterogeneity Among and Within Individuals." *American Economic Review* 103: 532–48.
- Gintis, Herbert. 2017. *Individuality and Entanglement: The Moral and Material Bases of Social Life*. Princeton, NJ: Princeton University Press.
- Gintis, Herbert. 2009. *The Bounds of Reason: Game Theory and the Unification of the Social Sciences*. Princeton, NJ: Princeton University Press.
- Gino, Francesca, Shahar Ayal, and Dan Ariely. 2013. "Self-Serving Altruism? The Lure of Unethical Actions that Benefit Others." *Journal of Economic Behavior and Organization* 93: 285–92.
- Gino, Francesca, Erin Krupka, and Roberto Weber. 2013. "License to Cheat: Voluntary Regulation and Ethical Behavior." *Management Science* 59: 2187–203.
- Gino, Francesca, Maurice E. Schweitzer, Nicole Schweitzer, and Dan Ariely. 2011. "Unable to Resist Temptation: How Self-Control Depletion Promotes Unethical Behavior." *Organizational Behavior and Human Decision Processes* 115: 191–203.
- Gneezy, Uri. 2005. "Deception: The Role of Consequences." *American Economic Review* 95: 384–94.
- Gneezy, Uri, Christina Gravert, Silvia Saccardo, and Franziska Tausch. 2017. "A Must Lie Situation: Avoiding Giving Negative Feedback." *Games and Economic Behavior* 102: 445–54.
- Gneezy, Uri, Bettina Rockenbach, and Marta Serra-Garcia. 2013. "Measuring Lying Aversion." *Journal of Economic Behavior & Organization* 93: 293–300.
- Gneezy, Uri, Alexander Imas, and Kristóf Madarász. 2014. "Conscience Accounting: Emotion Dynamics and Social Behavior." *Management Science* 60: 2645–58.
- Gneezy, Uri, Agne Kajackaite, and Joel Sobel. 2018. "Lying Aversion and the Size of the Lie." *American Economic Review* 108: 419–53.
- Gneezy, Uri, and Aldo Rustichini. 2000. "A Fine Is a Price." *Journal of Legal Studies* 29: 1–17.
- Grolleau, Giles, Martin G. Kocher, and Angela Sutan. 2016. "Cheating and Loss Aversion: Do People Lie More to Avoid a Loss?" *Management Science* 62: 3428–38.
- Haley, Kevin J., and Daniel M.T. Fessler. 2005. "Nobody's Watching? Subtle Cues Affect Generosity in an Anonymous Economic Game." *Evolution and Human Behavior* 26: 245–56.

- Hallsworth, Michael, John A. List, Robert D. Metcalfe, and Ivo Vlaev. 2017. "The Behavioralist as Tax Collector: Using Natural Field Experiments to Enhance Tax Compliance." *Journal of Public Economics* 148: 14–31.
- Hanna, Rema, and Shing-Yi Wang. 2017. "Dishonesty and Selection into Public Service: Evidence from India." *American Economic Journal: Economic Policy* 9: 262–90.
- Herrmann, Benedikt, Christian Thoni, and Simon Gächter. 2008. "Antisocial Punishment Across Societies." *Science* 319: 1362–7.
- Houser, Daniel, John A. List, Marco Piovesan, Anya S. Samek, and Joachim Winter. 2016. "Dishonesty: From Parents to Children." *European Economic Review* 82: 242–54.
- Hugh-Jones, David. 2015. "Honesty and Beliefs about Honesty in 15 Countries." Mimeo. University of East Anglia.
- Kajackaite, Agne, and Uri Gneezy. 2017. "Incentives and Cheating." *Games and Economic Behavior* 102: 433–44.
- Kahneman, Daniel, and Amos Tversky. 1979. "Prospect Theory: An Analysis of Decision under Risk." *Econometrica* 47: 263–91.
- Khalmetski, Kiryl, and Dirk Sliwka. 2017. "Disguising Lies - Image Concerns and Partial Lying in Cheating Games." CESifo Working Paper No. 6347.
- Kocher, Martin G., and Simon Schudy. In press. "I Lie? We Lie! Why? Experimental Evidence on a Dishonesty Shift in Groups." *Management Science*.
- Leibbrandt, Andreas, Maitra Pushkar, and Ananta Neelim. 2017. "Large Stakes and Little Honesty? Experimental Evidence from a Developing Country." Monash Business School, Department of Economics Discussion Paper No. 13/17.
- Mazar, Nina, and Dan Ariely. 2006. "Dishonesty in Everyday Life and Its Policy Implications." *Journal of Public Policy and Marketing* 25: 117–26.
- Mazar, Nina, On Amir, and Dan Ariely. 2008. "The Dishonesty of Honest People: A Theory of Self-Concept Maintenance." *Journal of Marketing Research* 45: 633–44.
- Mead, Nicole L., Roy F. Baumeister, Francesca Gino, Maurice E. Schweitzer, and Dan Ariely. 2009. "Too Tired to Tell the Truth: Self-Control Resource Depletion and Dishonesty." *Journal of Experimental Social Psychology* 45: 594–7.
- Mullainathan, Sendhil, and Marianne Bertrand. 2004. "Are Emily and Greg More Employable than Lakisha and Jamal? A Field Experiment on Labor Market Discrimination." *American Economic Review* 94: 991–1013.
- Peer, Eyal, Alessandro Acquisti, and Shaul Shalvi. 2014. "‘I Cheated, but Only a Little’: Partial Confessions to Unethical Behavior." *Journal of Personality and Social Psychology* 106: 202–17.
- Ploner, Matteo, and Tobias Regner. 2013. "Self-image and moral balancing: An experimental analysis." *Journal of Economic Behavior & Organization* 93: 374–83.
- Pruckner, Gerald J., and Rupert Sausgruber. 2013. "Honesty on the Streets: A Field Study on Newspaper Purchasing." *Journal of the European Economic Association* 11: 661–79.
- Reuben, Ernesto, and Matt Stephenson. 2013. "Nobody Likes a Rat: On the Willingness to Report Lies and the Consequences Thereof." *Journal of Economic Behavior & Organization* 93: 384–91.
- Schindler, Simon, and Stefan Pfattheicher. 2017. "The Frame of the Game: Loss-Framing Increases Dishonest Behavior." *Journal of Experimental Social Psychology* 69: 172–7.
- Schweitzer, Maurice E., Lisa Ordonez, and Bambi Douma. 2004. "Goal setting as a motivator of unethical behavior." *Academy of Management Journal* 47: 422–32.
- Shalvi, Shaul. 2012. "Dishonestly Increasing the Likelihood of Winning." *Judgment Decision Making* 7: 292–303.

- Shalvi, Shaul, and David Leiser. 2013. "Moral Firmness." *Journal of Economic Behavior & Organization* 93: 400–7.
- Titmuss, Richard M. 1971. *The Gift Relationship: From Human Blood to Social Policy*. New York: Pantheon Books.
- Utikal, Verena, and Urs Fischbacher. 2013. "Disadvantageous Lies in Individual Decisions." *Journal of Economic Behavior and Organization* 85: 108–11.
- van der Weele, Joël J., Julija Kulisa, Michael Kosfeld, and Guido Friebe. 2014. "Resisting Moral Wiggle Room: How Robust Is Reciprocal Behavior?" *American Economic Journal: Microeconomics* 6: 256–64.

CHAPTER 10

ETHICS AND ECONOMICS

A Complex Systems Approach

JOHN B. DAVIS

10.1 INTRODUCTION: WHAT IS ETHICS AND ECONOMICS AS A SINGLE SUBJECT OF INVESTIGATION?

THE subject of this volume is ethics *and* economics. Accordingly, a fundamental issue it must address is: what makes “ethics *and* economics” a single subject of investigation. Indeed, ethics and economics taken in themselves each constitute distinct types of investigation, as reflected in the fact that they have long been carried out in two separate disciplines by two largely independent groups of researchers, respectively, philosophers and economists. Their long histories as distinct disciplines thus create a barrier to understanding how together they constitute a single subject of investigation, and the fact that they have each developed in relative independence from one another also suggests that there exist inherent barriers in the way each is understood to combining them in a single field of investigation. It seems a mistake, then, to proceed with a discussion of ethics *and* economics without first reaching an understanding of at least some of the issues involved. However, this does not seem to be the starting point for much work that exists on ethics and economics, which instead often investigates *how* ethics and economics are connected—for example, in discussions of equity-efficiency tradeoffs—without first establishing *what* the relationship between ethics and economics is.

Where should one start, then, if we are to identify ethics *and* economics as a single subject of investigation? Given their independent histories, it would be natural to begin by saying that because ethics and economics have long been essentially independent fields of investigation, understanding what ethics *and* economics is should be framed in terms of an established discourse on the relationships between different disciplines. For example, perhaps ethics *and* economics should be understood as an “interdisciplinary”

type of investigation that somehow falls between and yet also combines ethics and economics. However, this strategy encounters the problem that there exists more than one conception of how different disciplines are related, interdisciplinarity being only one such type of relationship, and considerable debate over the different types of relationships between different disciplines. One taxonomy of the possibilities is as follows:

Interdisciplinary research or collaboration creates a new discipline or project, such as interfield research, often leaving the existence of the original ones intact. *Multidisciplinary* work involves the juxtaposition of the treatments and aims of the different disciplines involved in addressing a common problem. *Crossdisciplinary* work involves borrowing resources from one discipline to serve the aims of a project in another. *Transdisciplinary* work is a synthetic creation that encompasses work from different disciplines. (Cat 2017: sect. 3.3)

So, it seems that before identifying what ethics *and* economics is, we need to start at an even more basic level, first examine more broadly the ways in which different systems of thinking can be relatively independent but also interact, and then use any understanding that this provides to make judgments about what kind of relationship exists between ethics and economics.

The strategy for doing this adopted in this chapter is closest to the transdisciplinarity idea in that it employs a complex systems approach that treats ethics and economics as two different systems that interact in a single complex system. A complex systems approach takes the interaction of different systems as its starting point, and perhaps paradoxically explains their relative independence in terms of how they interact. That is, it does not begin with different disciplines, set them out first independently of one another, and then ask how they interact, since this strategy runs the risk of assuming they are largely unaffected by their interaction. Let me illustrate with a comparison of two ways of thinking about virtue in relation to economics.

One might then begin with virtue ethics as investigated in ethics, and apply that virtue ethics in an entirely new setting, say, in circumstances where economic behavior appears to resemble virtue behavior, and then conclude that the essential logic of that virtue ethics is unaffected by its application to the economic world. The context of application, that is, does not matter because it has been assumed away at the start by an independent disciplines-based strategy of analysis. The conviction that ethics *and* economics is an interdisciplinary subject of investigation, as characterized above, then, would support this mode of analysis.

Alternatively, in a complex systems approach, it might instead be argued that “virtue-like” behavior in the economic world differs significantly from what a traditional virtue ethics assumes, because a change in context has transformative effects upon that behavior. That is, context of application matters. Standard economics assumes *Homo economicus* agents are rational optimizers, but rational optimization is not part of what virtue ethics traditionally involves, so explaining “virtue-like” behavior in the economic world transforms ethics’ conception of virtue ethics. This mode of analysis has transdisciplinarity as its point of entry.

My point here concerns our mode of analysis of ethics *and* economics. Specifically, if we believe that ethics *and* economics is a single subject of investigation, we need to adopt an approach that treats it as a single complex system. It may well be, of course, that the interaction between ethics and economics which an ethics *and* economics would investigate in many cases does not have transformative effects on either. For example, it could well be the case that “virtue-like” behavior in the economic world does not differ significantly from what a traditional virtue ethics assumes. A complex systems approach should allow for this “small” effects of interaction possibility just as it allows for a “large” effects of interaction possibility. Indeed, which results we get will likely depend on the issues investigated. Thus, to put this complex systems approach to work we need to determine what issues in ethics *and* economics deserve examination.

Rather than proceed in a piecemeal, issue-by-issue manner, in this chapter I will instead look more broadly at transformative effects that ethics and economics might each have on one another in terms of two opposing visions of what ethics *and* economics might be about: one in which ethics as a whole has transformative effects on economics as a whole—which I associate with a social economic vision of the subject—and one in which economics as a whole has transformative effects on ethics as a whole—which I associate with the standard or mainstream economics vision of the subject.

One advantage of beginning in this way, then, is that it raises fundamental issues about what ethics *and* economics is as a single subject of investigation. Is it largely framed by ethics, by economics, or by both more or less equally? Or better, assuming there are differences of opinion on this question among ethics *and* economics researchers, what does each position imply about what ethics *and* economics is?

My own view is that a social economic vision best describes ethics *and* economics. The mainstream economics vision, however, is likely more familiar to people because it has been so influential. Its individualist emphasis frames ethical issues in the language of economics. Thus, I close these introductory remarks with a brief description of social economics as a distinct approach within economics.

Social economics originated in eighteenth-century Europe as an investigation of cooperative economic relationships rooted in community social values—its object being the *économie sociale* or the *Gemeinwirtschaft*. Its main assumption is that human society is everywhere and always value-laden, and that explanations of economic life cannot be value-free. Market values, then, are not just prices at which goods are exchanged, but expressions of a community’s values regarding how social relationships ought to govern exchanges between people. Thus, what explains the dynamics of an economy is how social values and economic activity interact and evolve together. This means economics cannot be a purely positive science in the manner of the natural sciences, nor should markets and economic activity be explained in a mechanical, ahistorical way.

Contemporary social economics maintains these ideas, but understands the idea of community in a way appropriate to a globalized world. Communities are still local in regard to our most immediate and familiar social relationships, but communities also exist between people in more extended and diverse kinds of social relationships. In both

cases, then, those relationships are framed by social values which guide us in how we interact. Different social values are likely to be operative in different types of social economic interaction. Social economics investigates these differences to explain how economies work.

Let me, then, summarize the organization of this chapter. Section 10.2 begins by first providing a brief overview of the complex systems approach that the chapter employs, and then uses this to distinguish the main commitments underlying the competing social economic and the mainstream economics visions of ethics *and* economics. My goal here is to set up an interpretive framework for each vision that allows us to evaluate the key concepts and principles their respective views maintain regarding how ethics and economics relate to one another and create a single subject. For example, I will explain why they differ regarding whether economics should be thought value-laden or should be value-free in terms of how they differ regarding their different understandings of the interaction of ethics and economics.

Sections 10.3 and 10.4 inventory and discuss the key concepts and principles for each vision. Because mainstream economics' vision is the more influential of the two, I discuss it first in Section 10.3 and then discuss the social economic vision in Section 10.4. The latter vision, as less familiar and as often thought more controversial, is clearer when contrasted to the former. In both sections, I distinguish strong and weak versions of these respective views to show a range of opinion on the positions and issues discussed. Section 10.5 returns to the interpretive framework from Section 10.2 that underlies the chapter, and briefly discusses the future prospects for ethics *and* economics as a single subject of investigation.

10.2 A COMPLEX SYSTEMS APPROACH TO ETHICS AND ECONOMICS: TWO VISIONS OF THE SUBJECT

Of course, there are many views of what complex systems are, so ironically one needs to simplify what we say about complex systems in order to be able to apply this framework effectively. Accordingly, I draw on one, especially influential, straightforward conception of what a complex system is that was developed by Herbert Simon, who though well known as a Nobel laureate economist was also trained in political science, and was thus well aware that investigations between established disciplines might produce new subjects of investigation.¹

¹ These two areas of expertise hardly capture Simon's polymath capacities, which included computational science, political science, cognitive psychology, sociology, public administration, management studies, and philosophy of science. His complex systems thinking was arguably the product of his wide-ranging expertise.

Simon's (1962) main intuition about complex systems is that they are made up of multiple, relatively independent subsystems whose interaction

- influences both their own activities but also the performance of the whole complex system which taken together they make up,
- produces changes which in turn reverberate back upon how these subsystems interact,
- so that the interacting subsystems and the whole system constantly influence each other in a continuing dynamic of change.

Thus, Simon presupposes that interaction of relatively independent subsystems determines both the nature of complex systems as single entities and how they change over time.

Applying this conception to ethics *and* economics would then be to say that ethics and economics are each relatively independent subsystems, as their histories as separate disciplines implies, whose interaction explains ethics *and* economics as a single subject of investigation. Heuristically,

- a process of interaction between them “somehow” influences how each operates,
- this “somehow” determines what ethics *and* economics is,
- and this over-riding conception “somehow” feeds back upon and influences their interaction.

To conclude this section, I only explain the first “somehow” for each of the two visions of ethics *and* economics, and in the next two sections identify the additional “somehow” connections in relation to how the concepts and principles each vision employs produce their respective visions of ethics *and* economics. Throughout, I distinguish strong and weak thesis versions of each vision.

In the mainstream economic vision, then, markets are a fundamental type of human activity, and consequently non-market social processes, including ethical behavior, are either ultimately reducible to market processes (the strong thesis), or are relatively unimportant compared to market processes and should be set aside when they conflict with market processes (the weak thesis). Broadly speaking, the economy subsumes society, or *society is fully embedded in the economy*. Thus, in the interaction between social processes, here specifically ethical behavior, and market processes as two relatively independent subsystems, market processes, or economics, determine the scope and nature of ethical behavior, or ethics.

In contrast, in the social economic vision, social processes constitute the fundamental human activity and include many forms of behavior, of which market behavior is only one form. In some circumstances, then, social processes, including ethical behavior, preclude the existence of market processes and show the “moral limits of markets” (the strong thesis), and in other circumstances social processes simply limit or condition market processes and require “taming the market” (the weak thesis). Broadly speaking, society subsumes the economy, or *the economy is fully embedded within society*. In terms

Table 10.1 The mainstream and social economic visions of ethics *and* economics

Ethics <i>and</i> economics visions	Strong thesis	Weak thesis
Mainstream economics: society is embedded in the economy	Social processes including ethical ones are reducible to market processes	Social processes including ethical ones should be set aside when they conflict with market processes
Social economics: the economy is embedded in society	Social processes preclude market processes: "moral limits of markets"	Social processes condition market processes: "taming the market"

of the interaction between ethics and economics as two relatively independent subsystems, ethics determines the scope and nature of economics.

Table 10.1 summarizes the strong and weak versions of these two visions.

I turn in Sections 10.3 and 10.4 for each vision, to how each vision "somehow" determines what ethics *and* economics is, and how these over-riding conceptions of what it is "somehow" feed back upon and influence their interaction, including what this implies about economic and social policy. Section 10.5, then, closes the chapter with comments on the future status of ethics *and* economics as a single subject of investigation.

10.3 THE MAINSTREAM ECONOMIC VISION: SOCIETY EMBEDDED IN THE ECONOMY

If society is fully or essentially embedded in the economy, what does this tell us about what ethics *and* economics is as a distinct subject of investigation and how this conception influences the interaction between the two? The natural point of entry for addressing these questions is the mainstream position on the positive-normative distinction, which frames its interpretation of ethics *and* economics. Section 10.3.1 sets out this position. Section 10.3.2 then discusses efficiency as the main "normative" concept the mainstream employs. Sections 10.3.3 and 10.3.4 discuss two key principles of analysis central to mainstream reasoning about economic policy that bear on the relationship between ethics and economics: the externalities concept and cost-benefit analysis.

10.3.1 The Mainstream's Understanding of the Positive-Normative Distinction

Mainstream economics' understanding of the positive-normative distinction dates back to David Hume's strong separation of the positive and normative based on a distinction in

language between “is” statements and “ought” statements. “Is” statements, Hume argued, concern facts, and “ought” statements concern values. Lionel Robbins (1935) made this distinction fundamental to mainstream (then neoclassical) economics when he inferred that it implied a distinction in types of discourses, associating science exclusively with positive statements and ethics primarily with normative statements. What does this view, then, imply about economics or science and ethics respectively?

Hume had argued in his famous “is-ought” doctrine (the “Humean guillotine”) that “ought” type statements can never be inferred from “is” type statements (1739: 469–70), from which it followed that science and ethics should function as entirely separate domains with science being the domain of positive statements and ethics being the domain of normative statements. The risk of confusing the two has serious consequences for science, which is undermined when “ought” type statements are admitted, and should statements of values replace statements of fact. In contrast, ethics, particularly in the form of social policy, should be grounded in “is” type statements, so in principle ethics is strengthened by contact with science. This leaves open the status of values, and mainstream economics has both strong and weak thesis positions on the matter.

The strong thesis, dating again back to Robbins, is that values are highly subjective and thus often arbitrary, whereas facts are intersubjective and objective. Science thus needs to distance itself from ethics. The weak thesis is that values need not always be subjective and arbitrary, though why this might be the case is usually left unexplored and secondary to the importance invested in science as a discourse that should be essentially value-free. An exception allowed in both cases is that methodological value judgments, which concern how science is practiced—such things as the desirability of clarity, consistency, a reliance on evidence, and so forth—are thought to have no effect on economics’ nature as a positive science (cf. Blaug 1998).

Overall, then, the mainstream understanding of ethics *and* economics, based on its understanding of the positive-normative distinction, does two things. First, it treats ethics and economics quite asymmetrically. Economics and science are highly valued and ethics is seen as at best an unreliable domain of intellectual investigation which is separate from economics. Second, the mainstream view is quite unclear about the scope and nature of values since it proscribes ethical ones, openly adopts methodological ones, and says little about what justifies this difference, while still treating economics as if it were value-free. Not surprisingly, this often leads to the charge that mainstream economics is positivist in comparison with other social sciences.²

If we emphasize just the asymmetry point, then, the mainstream vision fits the weak thesis above: namely, the idea that markets are a fundamental type of human activity, and that other social processes including ethics are relatively unimportant compared to market processes and can accordingly be largely ignored or left to others. However, if we

² In fact, Robbins was influenced by interwar logical positivism, which followed Hume in the strong separation of facts and values and subjectivist view of the latter as associated with emotivism, the meta-ethical doctrine that values are merely expressions of feeling and attitude (cf. Davis 2015).

emphasize the mainstream's positivism and ambiguity about the scope and nature of values that allows for methodological values as if they are not values while proscribing ethical ones, the mainstream vision can be seen to better fit the strong thesis that denies ethics' coherence, or even reduces ethical values to market values.

The strong thesis may seem improbable, especially compared to the more tolerant weak thesis, but in the next subsection I argue that there exists a basis for it in the standard understanding of the only "normative" concept used in mainstream economics: namely, the efficiency concept.

10.3.2 Efficiency as a "Normative" Concept

Efficiency, or Pareto optimality, is the idea that states of affairs which make improvements in how competitive markets work that leave no one worse off, as reflected in their given preferences, ought to always be promoted. Thus, on the surface the efficiency concept is explicitly normative since its expression requires an "ought" type statement. Yet at the same time, as a normative concept, efficiency as understood by most mainstream economists is different from virtually all other normative concepts. Let me emphasize two points. On the one hand, that an efficient state of affairs leaves no one worse off in preference satisfaction terms means that from everyone's own point of view such states of affairs should always be promoted. On the other hand, while the preference concept of course has a technical meaning in economics associated with the standard axioms regarding preferences that apply to it, one popular interpretation of it is as a comprehensive measure by which people fully evaluate any and all of their choice opportunities, as in Daniel Hausman's (2012) characterization of preference as providing a "total subjective evaluative comparison" of all things that might enter into people's choices.

What these two points suggest is that efficiency judgments are essentially indisputable. If a state of affairs *A* is efficient relative to a state of affairs *B*, then there exist no grounds for promoting *B* and not promoting *A*. The idea is that if preference evaluations are comprehensive of everything, not only can efficiency judgments never be reasonably contested, but should people appear to have other normative concerns, the all-inclusive view of preferences implies that those normative concerns must be implicit in their preferences. An older view of preferences as essentially self-regarding precluded this argument, because then other normative values could then be set against efficiency as a product of self-regarding preferences. Yet mainstream economists have given up this older, narrower view of preferences, become steadfastly agnostic about their content, and thus removed the need to refer to other normative concepts in judging competing states of affairs.³

³ Thus, a variety of kinds of non-standard preferences are typically excluded from consideration, such as, for example, preferences expressing envy.

There are two opposing ways, then, to interpret this outcome. On the one hand, it can be argued that the full range of ethical thinking that philosophers have long investigated can now be seen as implicit within mainstream economics explanation of choice behavior. To be clear, this does not mean that economics as a science is no longer value-free, because it is not the science itself but people's own values as embedded in their preferences that might include other ethical values. On the other hand, it can also be argued that the efficiency concept makes any reference to other ethical values effectively redundant in economics. All that is needed, that is, for the analysis of competitive markets is the efficiency concept. Which interpretation, then, is more likely?

The first view, that the full range of ethical thinking philosophers have long investigated can be seen as implicit in economics, is pretty clearly at odds with economists' general disinterest in ethical values. In contrast, the second view that the efficiency concept is sufficient in economics reflects much of mainstream economics' ordinary practice, and so seems more likely. I argue, then, that the centrality of the efficiency concept in economics supports the reductionist strong thesis. Indeed, if we understand the efficiency concept as simply a positive description of a particular state of affairs—namely, one in which as a matter of fact no person is worse off and at least one better off—one does not need to make any reference to it as an ethical idea at all. It is simply a positive description of a certain state of affairs without including any recommendation that it be achieved, or any analysis of its normative desirability. Economists may happen to believe that Pareto optima are normatively desirable, but this is incidental to their positive analysis of markets. For economics, then, ethical values are fully reducible to market values, not, as in the weak thesis, simply a relatively unimportant independent domain with which economics interacts.

10.3.3 Externalities

Addressing externalities is fundamental to mainstream economic reasoning about markets. Market transactions are assumed to primarily affect only those engaged in them, but markets are not perfect, and often give rise to unintended, third-party effects on others “external” to those transactions: spillover effects or externalities. Further, while (free and competitive) market transactions where externalities do not occur must be efficient, because the parties to those transactions see themselves as better off and no one is worse off, market transactions where they do occur can be inefficient because while the parties to those transactions are better off, third parties can be worse off. Thus, economic analysis of externalities and how markets function is necessarily framed in terms of realizing efficiency. This is often understood to be a matter of “internalizing” externalities, or assuring that spillover costs are borne by the transacting parties, not by third parties.

Thus, I suggest that the strong thesis emphasizes “internalizing” externalities, and frames “externalities”—despite the sense of the term—as phenomena properly seen as internal to market processes once markets are made efficient and function competitively. Then, if efficiency is seen as essentially a positive, non-normative concept, externalities

are basically understood as disturbances in markets rather than as reflecting normative issues. In effect, there is nothing truly “external” to market processes. This, then, is consistent with the strong thesis regarding ethics *and* economics and its reductionist approach to ethics.

In contrast, the weak thesis allows that there are circumstances where market transactions have *ethical* spillover effects, such as when people regard market transactions as unfair or unjust, even if they are otherwise efficient. An example might be low wage employment. In this case, rather than deny that such effects occur in a competitive market system, emphasis falls on how they should be treated. Thus, for negative externalities such effects are regarded as a special type of cost that must somehow be borne by the parties to the relevant market transaction. However, this translates an ethical judgment about right and wrong into a question of money expenditure, in this case puts a price tag on fairness or justice, and makes “internalizing” the externality a matter of how market processes determine the meaning and scope of ethical ones.

Thus, the strong thesis regarding externalities eliminates ethics and the weak thesis translates it into the language and metrics of economics. Consider now the nature and role of cost-benefit analysis in mainstream economics.

10.3.4 Cost-Benefit Analysis

While the efficiency concept is fundamental to mainstream reasoning, in practice economists often find it necessary to evaluate circumstances in which some state of affairs *A* is recommended over another *B* even though some individuals are made worse off in *A*. They then reason in cost-benefit terms, which involves showing that the balance of benefits and costs associated with *A* is superior to that associated with *B*. This raises an interesting question: if efficiency is no longer the basis for policy recommendation, and the seemingly broad idea of “benefit” has come into play, might normative principles regarding what is good then play a role in this less restricted form of reasoning, thus opening up economics to a more substantial engagement with ethics?

One reason to think that the idea of “benefit” might have broad normative meaning in standard economics is that individuals determine what benefits them according to their preferences, and as we saw above many economists believe that preferences need not be narrow and self-regarding, but may apply to anything they might value (Hausman 2012). Accordingly, cost-benefit analysis must weigh whatever people value against costs, whether it is payoffs to themselves or the good of others, and this accordingly opens the door to the possibility that a whole range of ethical views regarding what is good drives people’s preferences and might thus figure in an “ethically informed” cost-benefit analysis.

Yet the standard view in economics regarding preferences, while potentially open in regard to the scope of preferences, remains agnostic in regard to how they are factored into cost-benefit analysis. This tells us what would be a strong thesis in the mainstream regarding cost-benefit analysis: namely, that the content of preferences is irrelevant to

registering people's perceived benefit. Economists, that is, may ignore any ethics-related dimensions of what people register as a likely benefit, and simply through some preference elicitation method construct a value to put into the balance with expected costs. It may then be shown that a state of affairs *A* ranks higher than a state of affairs *B*, but it is irrelevant whether people believe the benefit from *A* is associated with greater happiness, being able to lead "good" lives, their social relationships being just, and so on. Cost-benefit analysis seen in this way does not require we discriminate between different views of what is ethically good, and indeed ignores altogether any meanings people associate with "good."

The weak thesis regarding cost-benefit analysis mirrors the weak thesis regarding externalities. The point of cost-benefit analysis is to reach decisions for ranking alternative possible states of affairs. Different people's views of why a certain state of affairs is good may be helpful in creating indices useful for producing aggregate measures of good. Suppose one group of people report that they prefer a certain state of affairs because it makes them happier, and another group reports that they prefer it because they believe it is more just. All the economist needs to do is provide weights for these different views (perhaps according to their frequencies in a population), and then generate a single measure of the "good" that state of affairs would produce. Thus, again, the meaning of what is good—a chief concern of normative reasoning—is collapsed into a market money metric.

The differences between the strong and weak views in regard to cost-benefit analysis are modest, and in both cases cost-benefit analysis ends up being essentially as restrictive as efficiency reasoning and standard analysis of externalities. In general, across all three subjects discussed in this section, the difference between the strong and weak views is a difference between eliminating ethics from economics and translating ethics into economics. Mainstream economics' view of ethics *and* economics, then, is that it basically can be explained fully in terms of economics. This reflects its guiding conception that society is embedded in the economy and social processes are explainable as market processes.

Table 10.2 summarizes the results of this section.

Table 10.2 The mainstream economic vision: Society embedded in the economy

	Strong thesis: reductionism	Weak thesis: exclusion and translation
Positive-normative distinction	Science is objective and ethics is subjective	Ethics unimportant compared to science
Efficiency concept	Positive, non-normative	Other normative concepts are not needed
Externalities with ethical aspects	Internalized in competitive markets	Treated as "costs"
Cost-benefit analysis	Benefit is only a measure of preference	Good is translated into monetary terms

10.4 THE SOCIAL ECONOMIC VISION: THE ECONOMY EMBEDDED IN SOCIETY

In the social economics vision of what ethics *and* economics is, the economy is embedded in society and social processes determine the nature and scope of market processes. To examine what this involves I begin with the same starting point used in my discussion of mainstream economics' vision, namely, how a social economic approach understands the positive-normative distinction. Section 10.4.1 sets out this view, and then distinguishes strong and weak theses regarding how ethics and economics are related to one another as separate yet "entangled" subjects. Section 10.4.2 then turns to what the social economic vision treats as alternative to the mainstream's emphasis on efficiency: namely, a much broader view of how normative principles pervade economic life. To structure this view, I frame it in terms of four different types of capabilities and moral values associated with those capabilities, and then also distinguish strong and weak views on the matter. Section 10.4.3 then returns to the issue of "entanglement" and (parallel to Sections 10.3.3 and 3.4) discusses two guiding principles regarding social policies towards markets: prohibition and intervention.

10.4.1 Social Economics' Understanding of the Positive-Normative Distinction

Whereas the mainstream's view of the positive-normative distinction is based on Hume's "is-ought" distinction, the social economic view is that this involves a mistaken understanding of the nature of language. Of course, "is" statements can be distinguished from "ought" statements when different statements contain only those terms. However, the mainstream inference that a statement free of "ought" language or its equivalents must be purely descriptive and value-free is at odds with modern philosophy of language's emphasis on the determinants of meaning. Indeed, it is generally argued by philosophers today that the meanings of terms are highly dependent on context and how they are used in speech and communication. A consequence of this is that terms that appear to be entirely descriptive can also convey normative and prescriptive content depending on how people interpret and use the statements in which they appear. It follows that descriptive statements may well be (though are not necessarily) value-laden. It further follows that we cannot assume that science is value-free simply because it is expressed in descriptive language. This applies especially to economics and the social sciences where the terms used in theories have multiple meanings and layers of interpretations.

An influential recent example of this understanding is Hilary Putnam's characterization of the descriptive term "cruel" used in reference to infamous Roman emperors. The statement "Nero was a cruel Roman emperor" is descriptive, but the term "cruel" also implies for most people that we ought to judge Nero as inhumane and immoral

(Putnam 2002: 34; cf. Davis 2015). Putnam labels this mixing up of descriptions and prescriptions and facts and values in ordinary language the *entanglement thesis*. Contrary to Hume, it allows us to infer “ought” statements from “is” statements. Nonetheless, this does not imply that we cannot distinguish science and ethics. Clearly, they are different types of investigations. The point rather is that different types of investigations can influence and affect one another: a symmetric or two-way street relationship in the case of ethics and economics, rather than the asymmetric one-way street from economics to ethics that the mainstream assumes.⁴

Consider in the case of economics, then, the status of the theory of rational choice. If one reviews textbook accounts, one finds the theory explained in a purely descriptive way as if it were value-free. Yet the history and development of the theory is bound up with—entangled with, Putnam would say—defenses of individual freedom and autonomy, which are normative ideals. Certainly, the theory retains a descriptive function, and can be used to describe choice behavior without making reference to those ideals. However, that does not mean that those associations and meanings are absent, so, as in Putnam’s Nero example, rational choice theory is looked upon by economists favorably both for its descriptive usefulness and also because it implicitly prescribes these normative ideals. Thus, rational choice theory is value-laden.

This then raises the question: how are ethics and economics as two types of discourse related to one another in the social economic vision when they are both separate domains of investigation and yet also entangled? At the end of Section 10.2, I distinguished a strong thesis and a weak thesis for the social economic vision that market processes are embedded in social processes. Let me state them here more fully.

The strong thesis—which I associate especially with the “moral limits of markets” view—responds to the entanglement idea by distinguishing a domain of social life from which market processes should be excluded. Its argument is essentially that normative principles have meanings that cannot be reduced to the logic of market processes, that doing so is morally repugnant (Roth 2007) and tends to “crowd out” those principles, so that there are certain domains of social life that should simply be “off limits” to markets. That is, some things are “priceless.”

I distinguish two versions of the strong thesis. On the one hand, one might emphasize that markets are often destructive of normative principles, as Michael Sandel (2012) has argued in connection with the idea that there are some things that “money can’t buy.” For example, although insurance markets and court proceedings put a price on people’s lives, Sandel argues that human life cannot have a price. This focus, however, sets aside the issue of whether ethics is entangled with economics within markets themselves. On the other hand, then, a second version of the strong thesis argues that different ethical principles operate within markets and outside markets, so that the moral limits of markets view is rather a matter of determining which belongs in which. For example, one might argue that even if market processes exhibit commutative justice (the idea that an exchange of equals is equitable) they cannot be expected to address distributive justice

⁴ On this, see Elgin’s (1989) discussion of the relativity of facts and objectivity of values.

(the socially equitable distributions of goods). Both versions of the strong, “off limits” thesis, then, emphasize establishing normative boundaries on market activity.

The weak thesis—what I term the “taming the market” view—focuses on how markets need to be transformed to be ethically acceptable, less than on what is normatively “off limits” to markets. This perspective emphasizes policies designed to change how markets themselves function, and aims at making ethical values central to their performance and the effects they have on market prices and quantities. In effect, the strategy is to replace (or augment) the mainstream’s sole reliance on efficiency reasoning as a means of regulating markets. Indeed, to the extent that efficiency reasoning is thought by some mainstream economists to be a non-normative, value-free type of intervention in markets, the conviction behind the “taming the market” view is that markets are thoroughly value-laden, and the issue is then which values ought to prevail. Below I will identify two different types of strategies this approach involves: ones aimed at directly intervening in how market outcomes are produced—a direct strategy—and ones aimed at changing the underlying conditions affecting how market outcomes are produced—an indirect strategy.

Given this understanding of the positive-normative distinction and these two theses, I turn to a social economic understanding of how normative principles pervade economic life.

10.4.2 Capabilities, Moral Values, and Human Dignity

If, contrary to mainstream economics the positive and normative are not strongly separated, then we need to explain how economic life is value-laden. There are a variety of ways that this can be done, but one particularly influential way involves the capability approach developed by Amartya Sen (1987), who believes one of the main failures of modern economics is its tendency to become an “engineering” science and consequent failure to integrate ethics and economics. As a means of correcting this, his capability approach builds economics around a broad conception of freedom that concerns people’s activities in all domains of life. Thus, in referring to people’s capabilities, Sen states, “What people can positively achieve . . . is influenced by economic opportunities, political liberties, social powers, and the enabling conditions of good health, basic education, and the encouragement and cultivation of initiatives” (1999, 5).

Sen then makes two distinctions that he uses to distinguish four different forms of what may be to a person’s individual advantage. The first distinction is between what promotes a person’s well-being and what promotes a person’s overall agency goals, or “goals other than the advancement of his or her well-being” (1993: 35). The second distinction is between a person being able to actually achieve something and the person simply having the freedom to pursue the objectives she wants to achieve (*ibid.*). Combining these two distinctions, he labels the four forms of individual advantage people have as: well-being achievement, agency achievement, well-being freedom, and agency freedom.

I argue, then, that each of these four cases can be associated with a distinct moral value, so that promoting these capabilities is also a matter of promoting these different moral values.⁵ Specifically:

Well-being achievement concerns what all people should have, irrespective of their individual circumstances. Sen's own examples are being able to avoid premature mortality and being adequately nourished. Well-being achievement, then, is naturally associated with promoting equality since no one should be excluded from this most basic level of well-being.

Well-being freedom concerns what a person freely does to promote her well-being, even if she is not successful in achieving it. An example is the pursuit of higher education. In this case, it falls upon the person to take responsibility for promoting their own well-being, so the idea of personal responsibility is central.

Agency achievement addresses goals that people want to achieve which are not associated with their own well-being. For example, a disabled individual might value personal independence, even if they are not made better off by achieving it. When we place a moral value on this goal, we value a person's right to personal independence in itself. As something that should be achieved, it may be considered a right, and more generally as applies to all people who wish to have it, a human right.

Agency freedom also concerns goals other than personal well-being, but in this case, it is not their achievement that is important but rather a person's ability to pursue them. For example, people provide care for others, the elderly, refugees, the sick, and so forth, and this is done for its own sake whether or not improves others' well-being. The moral value that this promotes is freedom, in both its negative and positive senses, since care for others enhances their freedoms by both reducing constraints on their lives and increasing their power to pursue their lives.

Table 10.3 presents these four forms of individual advantage and the moral values they involve. To be clear, others might link moral values to these four cases differently from the ones proposed here, which are only offered as reasonable interpretations of what they each normatively imply. The main point is that economic life is value-laden,

Table 10.3 Four different types of individual advantage and associated moral values		
	Well-being	Overall agency goals
Achievement	<i>Well-being achievement: equality</i>	<i>Agency achievement: human rights</i>
Freedom to achieve	<i>Well-being freedom: Responsibility</i>	<i>Agency freedom: Negative and positive freedom</i>

⁵ Similarly, Davis and McMaster (2017) use Sen's framework to identify and link these four forms of individual advantage with particular moral values important to people's health care capabilities.

and that our examination of it needs to make this explicit, not proceed as if the positive and normative are strongly separated.

This analysis, then, tells us what sorts of moral values may be associated with different forms of economic behavior according to what people's aims are, but it does not tell us anything about the relative importance of the two in economic life. One might, for example, argue that most people are as a matter of fact mostly concerned with their well-being achievement and that, say, agency freedom, the furthest removed from that, rarely concerns them. This would imply that equality is especially important in economic life, and that securing others' negative and positive freedom is less important. Or, one might argue that all four forms are equally present in economic life, implying that all four types of moral values are as well. Other possibilities could also be considered.

Consequently, if we are to explain both *how* pervasive normative principles are in economic life and *what* normative principles tend to prevail in economic life, we need to be able to say something about the prevalence of these different forms of individual advantage in economic life themselves. That is, our "positive" economic analysis needs to guide us in establishing the scope and nature of our normative analysis in economic life. To proceed without this foundation and point of entry is to run the risk of simply imposing the analyst's moral preferences on one's explanation of economic life, irrespective of whether people's behavior justifies it—thus re-separating ethics and economics, but from the vantage point of ethics!

Thus, to broadly distinguish two social economic views of ethics *and* economics, here I distinguish strong and weak views regarding *how* value-laden economic life is and regarding *what* normative principles prevail in economic life, as associated with two positions on the importance of these different forms of individual advantage in economic life.

The weak thesis essentially assumes that well-being achievement and well-being freedom are the dominant concerns of most people in economic life, reflecting economics' long preoccupation with well-being. It follows that the moral values that most importantly underlie economic life are equality and responsibility. In effect, people ought to have equal chances when it comes to assuring basic education and health, and then it falls upon them individually and becomes their responsibility should they aspire to further levels of achievement. I characterize this as a weak thesis because of its more limited scope regarding *what* normative principles are important in economic life.

The strong thesis, then, can be contrasted with the weak thesis according to the greater scope it assumes applies to *what* normative principles are important in economic life. Reflecting Sen's influence in making goals other than well-being a part of ethics *and* economics, this view assumes that all four forms of individual advantage are important in economic life, and accordingly that, in addition to equality and responsibility, human rights and positive and negative freedom are important moral values in economics as well. Thus, for the strong thesis, the economy and economics is highly value-laden, whereas for the weak thesis they are modestly value-laden.

Table 10.4 sets out these two positions.

Table 10.4 Positions regarding the pervasiveness of moral values in economic life

Positions	Forms of individual advantage	Moral values
<i>Weak thesis</i>	Well-being achievement and well-being freedom	Equality and responsibility
<i>Strong thesis</i>	Well-being achievement, well-being freedom, agency achievement, agency freedom	Equality, responsibility, human rights, negative and positive freedom

I close this section, then, with a brief comment regarding human dignity as the over-arching value in the social economic approach to ethics *and* economics. The capability approach supports multiple social values and can be interpreted in weak and strong thesis terms. This suggests that it lacks a unifying vision, so that how moral values enter into economic life has no clear interpretation. However, what this approach emphasizes overall, I argue, is the moral value of the person, or the central importance of human dignity. To see this, consider how the entanglement view treats individuals as compared to the mainstream economic conception of individuals.

In light of its strong separation of the positive and normative, the mainstream conception explains individuals in purely positive terms in that their behavior depends on their private preferences, which being subjective in nature have no particular relationship to social values. In contrast, in an entangled world individual behavior cannot be explained without reference to values, and as a consequence individuals are embedded in a world of social values. It follows that individuals are moral beings, and as such are themselves objects of moral concern. What does it mean, then, to say that individuals are objects of moral concern? In the social economic vision, it is to invest individuals with moral importance, or regard individuals as entitled to moral respect. This is to invest individuals with human dignity, and, assuming that individuals are fundamental to the social world, it makes human dignity an over-arching moral value underlying the social economic approach to ethics *and* economics.

In the next two subsections, then, I turn to two approaches to social policies central to the social economic vision of ethics *and* economics.

10.4.3 The Moral Limits of Markets View: Prohibitions-Based Social Policies

I characterized the moral limits of markets view as the entanglement strong thesis, and distinguished two versions. Sandel's version emphasizes what "money can't buy" or what should not be priced in certain domains of social life—what Roth treats as morally "repugnant"—and says that these domains are simply "off limits" to market processes. The other version of the strong thesis instead argues that different ethical principles

operate within markets and outside markets, so that the moral limits of markets view is rather a matter of determining which principles belong in which domains.

What these two versions share in social policy prescription terms is the idea of hard and fast prohibitions on where and how markets ought to operate. Social policy prescriptions reflect policies that are cast at the level of society. As such, they do not involve the sorts of policies that mainstream economics employs, where the focus is on adjusting the performance of markets through changes in prices and incomes. Rather, policy typically operates at an institutional level in adjusting the place and role of markets as a particular type (and only one type) of social institution. Needless to say, policy conceived in this way as social policy is foreign to the thinking of those who subscribe to the opposite view that society is fully embedded in the market process. At the same time, proponents of “off limits” thinking find the idea that social policies regarding markets are a matter of adjusting prices and incomes short-sighted and misguided.

What, then, are the main mechanisms used to establish the moral limits of markets? Prohibitions first and foremost take the form of laws that compel certain types of behavior irrespective of people’s economic incentives, thus demonstrating the reach and priority of law as a social domain independent of the market process. Not all laws are motivated by moral considerations, but many are, such as laws prohibiting owning people or slavery. Secondly, prohibitions take the form of custom and convention. Again, custom and convention is not always motivated by moral considerations, but they often are, as for example when it is said that elected representatives in democratic societies should not allow themselves to be influenced by personal gain.

10.4.4 The Taming-the-Market View: Interventions-Based Social Policies

I characterized the taming-the-market view as the entanglement weak thesis. Rather than seeking to segregate social domains according to the normative principles that apply or do not apply within them, the aim in this instance is to change the way markets function to make them normatively more acceptable. Markets are value-laden, then, but usually insufficiently so. While markets may exhibit efficiency, they should also, for example, function in an equitable way. This version of the entanglement thesis is thus weak compared to the strong thesis in that the degree of change it calls for through social policy is more modest. Nothing is necessarily “off limits” to markets; markets simply need to give further support to normative considerations in how they function. I then distinguish the policy strategies this view involves as direct and indirect in regard to how they aim to change the ways that markets function.

Specifically, the direct strategy aims at changing market outcomes by intervening in the way that markets adjust. One such strategy involves imposing price controls such as price floors in the case of minimum wages and price ceilings in the case of rent controls. By setting floors and ceilings, prices are prevented from moving to equilibrium levels.

Alternatively, the direct strategy can take the form of quantity controls, such as when air and water quality is protected by capping industrial effluents, the result of which again is that markets do not adjust to equilibrium levels. Thus, in both cases rather than only supply-and-demand forces, normative principles determine how markets adjust.

In contrast, the indirect strategy aims at changing market outcomes by influencing the underlying determinants of market behavior. For example, taxes and subsidies shift supply and demand curves, and this changes equilibrium prices and quantities. “Sin” taxes on goods believed to be harmful or undesirable raise their prices and discourage their consumption. Public spending for education lowers its price by increasing its supply. Market adjustment operates normally, but changing the basis on which goods are demanded and supplied improves the outcomes that result in a normative way.

Needless to say, both the direct and indirect intervention strategies can be used in a complementary way to achieve any given normative goal. Indeed, interventions-based social policies and prohibitions-based social policies above can also be used in a complementary way to achieve the same normative goals. What distinguishes the interventions approach, then, is the idea that since markets have significant effects on normative goals, they need to be made more value-laden, whereas the prohibitions approach sees markets themselves as less entangled. Thus, the two broadly different approaches here ultimately differ on their views of entanglement and value-ladenness.

10.5 FUTURE PROSPECTS FOR ETHICS AND ECONOMICS AS A SINGLE SUBJECT OF INVESTIGATION

The point of entry for this chapter, the status of ethics *and* economics, asks whether the two types of investigation involved function as a single subject of investigation or rather as an intersection of two largely independent disciplinary investigations. As Cat’s (2017) taxonomy of different forms of disciplinary interaction makes clear, there exist quite different ways in which we can understand how different disciplines interact, and consequently we need to proceed on a case-by-case basis. What this chapter did, then, was map how the two main visions of the relationship between ethics and economics, based on their opposed understandings of the positive-normative distinction, support two possible, quite different conceptions of ethics *and* economics.

The fundamental importance of the positive-normative distinction to these two conceptions reflects its deep philosophical nature. Not only are important philosophical questions about central concepts in human thinking involved, but far-reaching questions about the nature of science and intellectual investigation are part of how this distinction is understood. This might seem to imply, because deep philosophical issues are perennial, that the relationship between ethics and economics, to the extent that it depends on how this distinction is understood, will always remain unresolved, and accordingly we

will always have two competing visions of the relationship between these two kinds of investigation and two domains of social life.

This may well be what lies in the future for ethics *and* economics, but there is also a reason to think that the two visions will be replaced by a single one, specifically, the social economic vision. I indicated at the outset that, in my view, the social economic vision best describes the relationship between ethics *and* economics. In Davis (2015), I explained why I think this: essentially that the philosophical arguments against the Humean view of the positive-normative distinction that underlie the mainstream view are correct and that philosophers are in agreement that they are correct. Further, I argued that proponents in economics of the Humean view either ignore the philosophical arguments against it or suggest reasons for defending it that do not get exposed to serious scrutiny, either from philosophers or others. Thus, proponents of the Humean view seem to have little more to defend their position than loyalty to an inherited, unexamined thinking that has characterized mainstream economics since Robbins. This suggests that the foundation of their vision of economics as an essentially positive science is tenuous at best.

Of course, this does not tell us that the social economic vision of ethics *and* economics will ultimately prevail. Factors other than rational argument often determine how disciplines develop over time. At the same time, in the long run history of the sciences and intellectual investigation views that do not have the support of the large majority usually fail and are abandoned, and mainstream economists are isolated in social science in regard to this issue. Thus, it seems that the future prospects for an entangled ethics *and* economics as a single subject of investigation seem to be good.

ACKNOWLEDGMENTS

Thanks for very helpful comments on a previous version of this chapter go to Zohreh Emami, Wade Hands, George DeMartino, and Mark White.

REFERENCES

- Blaug, Mark. 1998. "The Positive-Normative Distinction." In John B. Davis, D. Wade Hands, and Uskali Mäki (eds.), *The Handbook of Economic Methodology* (Cheltenham, UK: Edward Elgar), pp. 370–4.
- Cat, Jordi. 2017. "The Unity of Science." In Edward N. Zalta (ed.), *Stanford Encyclopedia of Philosophy*, available at <https://plato.stanford.edu/archives/fall2017/entries/scientific-unity/>.
- Davis, John B. 2015. "Economists' Odd Stand on the Positive-Normative Distinction: A Behavioral Economics View." In George DeMartino and Deirdre McCloskey (eds.), *The Oxford Handbook of Professional Economic Ethics* (Oxford: Oxford University Press), pp. 200–18.
- Davis, John B., and Robert McMaster. 2017. *Health Care Economics*. London: Routledge.
- Elgin, Catherine. 1989. "The Relativity of Fact and the Objectivity of Value." In Michael Krausz (ed.), *Relativism: Interpretation and Confrontation* (Notre Dame, IN: University of Notre Dame Press), pp. 86–98.
- Hausman, Daniel M. 2012. *Preference, Value, Choice, and Welfare*. Cambridge: Cambridge University Press.

- Hume, David. 1739. *A Treatise of Human Nature*, 2nd ed., L. A. Selby-Bigge (ed.), rev. P. H. Nidditch. Oxford: Clarendon Press (1978 edition).
- Putnam, Hilary. 2002. *The Collapse of the Fact/Value Dichotomy*. Cambridge, MA: Harvard University Press.
- Robbins, Lionel. 1935. *An Essay on the Nature and Significance of Economic Science*, 2nd edn. London: Macmillan.
- Roth, Alvin E. 2007. "Repugnance as a Constraint on Markets." *Journal of Economic Perspectives* 21 (3): 37–58.
- Sandel, Michael. 2012. *What Money Can't Buy: The Moral Limits of Markets*. New York: Farrar, Straus and Giroux.
- Sen, Amartya. 1987. *On Ethics and Economics*. Oxford: Basil Blackwell.
- Sen, Amartya. 1993. "Capability and Well-being." In Martha Nussbaum and Amartya Sen (eds.), *The Quality of Life* (Oxford: Clarendon Press), pp. 30–53.
- Sen, Amartya. 1999. *Development as Freedom*. New York: Knopf.
- Simon, Herbert. 1962. "The Architecture of Complexity." *Proceedings of the American Philosophical Society* 106: 467–82.

CHAPTER 11

ECONOMICS AND ETHICS WITHIN THE AUSTRIAN SCHOOL OF ECONOMICS

PETER J. BOETTKE AND KAITLYN WOLTZ

11.1 INTRODUCTION

THE Austrian School of Economics, in the minds of most professional economists, is identified with the libertarian wing of the economics profession. This is true whether one reads a popular principles of economics book, such as David Colander's *Economics* (2000) or a popular trade book such as Ha-Joon Chang's *Economics: A User's Guide* (2014). There should be no doubt that self-identified Austrian economists often emphasize the power of the market mechanism to improve the welfare of the masses, and the tyranny of politics as it is said to impoverish the masses while benefiting only the powerful and privileged few. However, this conclusion is not evidence of ethical assumptions influencing analysis or ideology. Rather, it is a conclusion drawn from analysis. If this is a conclusion of economic science, it is not a conclusion widely shared by fellow economists. How can the Austrian economists make such claims? We argue that Austrian economists separate ethical discussions or assumptions from their economic analysis. They maintain strict adherence to value-free analysis through an emphasis on social cooperation, which allows them to maintain their objectivity with respect to individuals' ends. This combination allows Austrian economists to maintain their positions as value-free scientists while arguing that a free-market, capitalist system will best achieve peoples' diverse ends. This conclusion often prompts charges of politically motivated analysis despite Austrian economists' discipline as value-free scientists.

This identification of the Austrian School of Economics as libertarian is ironic for at least two historical reasons. First, in one of the earliest serious books to address the unwarranted intrusion of politics into modern scientific economics, Gunnar Myrdal wrote "in Austria, economics has never had direct political aims in spite of the close

connection of the Austrian marginal utility theory with utilitarian philosophy. The Austrians were preoccupied with value theory and never elaborated a detailed theory of welfare economics” (1929: 128). Second, the Austrian school of economics, in particular the writings of Ludwig von Mises, inspired Lionel Robbins’ critique of the efforts by Pigou and others to offer a detailed theory of welfare economics. Ludwig von Mises, the inspiration for Robbins’ critique, had completely absorbed Max Weber’s (1917) arguments for *wertfrei*, or value-free, social science. Weber’s argument was more than a rhetorical strategy in debate—it was a discipline for constructive conversations in scientific debate. It was an argument for positive analysis prior to the development of the positivist philosophy of science. Put simply, it argued that, as social scientists, we should not engage in arguments over the ends of various public policies, and instead restrict our analysis to the efficacy of the means chosen to pursue those given ends. By restricting the analysis to the logical examination of means and ends, the theorist can serve as the social critic without violating the intellectual discipline of *wertfrei* social science. To ignore these strictures, however, is to betray science and engage in normative philosophizing parading as positive analysis.

Following these strictures in his critique of socialism, Mises did not question the ends of socialism, but treated those ends as given. At the time, one of the critical ends of the socialist project was the rationalization of production.¹ Mises (1920) argued that the means chosen by the socialists—for example, collective ownership of the means of production—was incoherent with respect to the rationalization of production. Without private property in the means of production, economic decision makers would be unable to engage in rational economic calculation of alternative paths of investment and production. Absent the means of rational economic calculation, that rationalization of production would be impossible. Socialist means were incoherent with respect to socialist ends.

This is a decisive objection, and it is an objection that requires *no* assessment of values or political philosophy. It is an immanent critique of a social philosophical and political economy argument. Mises applied his *wertfrei* economic analysis not only to the extreme case of socialism, but also to less extreme examples of interventionism, such as price controls and the manipulation of money and credit by a central bank. Despite the caustic tone one can read in Mises, a careful reading of his work reveals that he strictly adhered to the discipline of the analysis of the efficacy of chosen means to given ends. He *never*, as an economist, questioned ends.

It was this discipline, with its depth of implication for important topics, that captured the imagination of Lionel Robbins in the 1920s. Robbins used what he learned from Mises to launch his own critical analysis based on positive economics against socialist and interventionist proposals in Britain, expose the pretensions of Pigou’s welfare economics, and challenge the emerging “new” economics of the Cambridge circus, which

¹ The goal of socialists was to bring society out of the “anarchy of production”—meaning the lack of central planning of production in capitalism—and instate the “rationalization of production”—meaning the central planning of production.

included Joan Robinson and Richard Kahn, around Keynes. In *An Essay on the Nature and Significance of Economic Science* (1932), Robbins elaborated the strictly value-free nature of economic science and demonstrated the pseudo-scientific nature of welfare economics which ultimately amounted to little more than normative philosophizing parading as positive analysis. As F.A. Hayek migrated from Vienna to LSE, he and Robbins represented a formidable challenge to the Cambridge hegemony in British economics at least throughout the 1930s. The Austrian school of economics, mixed with the contributions of Philip Wicksteed and Edwin Cannan, provided the foundations for a properly understood neoclassical economics of the LSE in the 1930s and 1940s. Subjectivism, processes, knowledge, and competition became the hallmark of this neoclassical approach. Economics was perceived as a purely scientific discipline that had profound implications. Economic science had put parameters on people's utopias.

Such parameters do not preclude prosperity; they encourage it. While the goal of the socialists was to engender a richer society from the virtuous cooperation of neighbors, Mises, Hayek, and Robbins demonstrated that their chosen means—collective ownership of the means of production—would fail. Rather than cooperating in achieving their different ends, neighbors would compete for resources. Socialism would foster fierce competition, not virtuous cooperation. Social cooperation is necessary for prosperity. Prosperity occurs when neighbors—even strangers—can hold different ends, but cooperate for the achievement of those different ends. Mises, Hayek, and Robbins emphasized that those rules which foster cooperation rather than competition are the key to prosperity and growth. Thus, they maintained their role as value-free scientists while arguing that free markets and capitalism would induce growth without sacrificing individuals' ends.

How ironic is this insistence on strict value-freedom in their writings, in light of the authors being the main intellectual opponents ideologically speaking against socialism and Keynesianism, as well as methodologically speaking against excessive aggregation, excessive formalism, and positivistic empiricism (Gunning 2005)? In fact, as "positive analysis" came to be understood by fellow economists *not* in terms of means-ends analysis, but, rather, as the willingness to subject hypotheses to empirical refutation, the Austrian school stance of *wertfrei* became less and less understood. As a result, the inference that the Austrian school of economics was simply a handmaiden to a more or less laissez-faire ideology became more prevalent among professional economists and in the popular imagination. This unfortunately has not changed. Mises and Hayek (if not Robbins) are better known for what they oppose than for their understanding of economic forces *at work* across a variety of social situations.

We welcome this opportunity, and in this context, to discuss once again, the relationship of economics and ethics from an Austrian school perspective. We will do so in the following manner. Section 11.2 will explore the question "Is economics value-free?" Section 11.3 will then explore further how the relationship between value-free analysis and ethical concerns is embedded in every question of the scale and scope of governmental action. Section 11.4 will focus on the emphasis of social cooperation in the Austrian school, and the insight into economic growth that an emphasis of social

cooperation provides. Finally, we conclude the chapter with summary remarks on the science of economics, the art of political economy, and ethical discourse.

11.2 IS ECONOMICS VALUE FREE?

Adam Smith's writings build on an argument that the right and the good are inexorably linked to each other. The "system of natural liberty" that he articulated also resulted in the "wealth of nations," so one can reasonably infer from reading him that what is right for mankind, is also good in terms of consequences for mankind. In other words, virtuous pursuits lead to prosperity. This is perhaps best captured in a conjecture made in lectures notes from 1755, 21 years before he published *An Inquiry into the Nature and Causes of the Wealth of Nations*. Smith said:

Little else is requisite to carry a state to the highest degree of opulence from the lowest barbarism, but peace, easy taxes, and a tolerable administration of justice; all the rest being brought about by the natural course of things. All governments which thwart this natural course, which force things into another channel, or which endeavour to arrest the progress of society at a particular point, are unnatural, and to support themselves are obliged to be oppressive and tyrannical.

(Quoted in Stewart 1793: 68)

In a presentation to the Royal Academy in Edinburgh in 1793, a few years after Smith's death, Dugald Stewart revealed these lecture notes he had in his possession to demonstrate the consistency of Smith's thinking through time. Smith's own system of thought has been subjected to interpretation and debate for centuries, with the most famous controversies surrounding the supposed tensions between *The Theory of Moral Sentiments* (1759) and *The Wealth of Nations* (1776). To many this shows the tension between social philosophy and political economy in general, while to others the reconciliation of the two texts demonstrates the intimate connection between political economy and social philosophy. We will not resolve this issue, but, for our purposes, the critical issue is the recognition that at the time of Smith, the system of natural liberty—what David Hume would describe as a system of property, contract, and consent—provided the institutional framework within which economic forces would work to lift mankind from abject poverty into material progress and human flourishing (Smith 1776: 51). The right and the good work together.

This basic conjecture spread among economic thinkers as the teachings of Hume and Smith were translated into French by Jean-Baptiste Say, and refined in the United Kingdom by David Ricardo and James Mill. But the French Revolution had a damning influence on such a simple conjecture about the right and the good. Claims of the "rights of man" no longer were simultaneously conjectures about "utility." As Jeremy Bentham famously said, rights claims were "nonsense upon stilts" (1843: 914), and utility is all that we can reasonably discuss. And, after Bentham, that is what political economists

and economists did, even to the chagrin of critics who vociferously proclaimed that economists know the price of everything but the value of nothing. In philosophical tracts, poems, and even novels, economists are portrayed as cold-hearted “philosophers of fact” without care or commitment to the betterment of mankind, let alone to the least advantaged among us.

How could it be that a systematic body of thought that began with an inquiry into the nature and causes of why some nations are rich while other nations are poor has come to be understood in such a manner? The narrative conjecture we will work with is that during the nineteenth century, advancements in the discipline of political economy demonstrated the correspondence between the aspiration of individual autonomy from oppression with improvements in material progress and generalized prosperity, but these same developments did not address the question of the distributive justice of the social system of exchange and production. Throughout the nineteenth century, the evidence continued to accumulate throughout Europe and North America about productive specialization and peaceful social cooperation that follows from the general institutional framework of property, contract, and consent. But within this experience, there was also a growing awareness of lingering social ills such as poverty, ignorance, and squalor. Urbanization highlighted these social ills. The wretched conditions of ill-fortune were more directly visible as neighbors were on top of neighbors. Rather than famine, the scourge of the modern age was mass unemployment due to industrial fluctuations.

This situation created a significant paradox: economics is vitally important for understanding the forces that define and shape the modern industrial world, but, as it was often claimed, lacked the tools necessary to fix the modern world. Some economic traditions, such as the Austrian school of economics in the late nineteenth century, were comfortable with the limiting role of “students of civilization” (Dekker 2016), but most practitioners wanted more out of economics. This was especially true of the Cambridge School from Alfred Marshall to John Maynard Keynes. In a 1908 lecture, “Economic Science in Relation to Practice,” A.C. Pigou summarizes the position as elegantly as we have seen:

If it were not for the hope that a scientific study of men’s social actions may lead, not necessarily or immediately, but at some point and in some way, to practical results in social improvement, I should myself – I do not intend to speak for others – I should myself regard the time devoted to that study as time misspent. That seems to me true of all social sciences, but especially true of Economics. For Economics “is a study of mankind in the ordinary business of life”; and it is not in the ordinary business of life that mankind is most interesting or inspiring. If I desired knowledge of man apart from the fruits of knowledge, I should seek it in the history of religious enthusiasm, of passion, of martyrdom and of love; I should not seek it in the market-place. When we elect to watch the play of motives that are ordinary in man – that are sometimes mean and dismal and ignoble in man – our impulse is not the philosopher’s impulse, knowledge for the sake of knowledge, but rather the physiologist’s, knowledge for the healing that knowledge in some measure may help to bring. (1908: 11)

This idea that economics must bear “fruit” in ethical discourse and in improving the conditions of man has been the subject of intense discussion from Marshall to Joseph Stiglitz, who in his work *Whither Socialism?* (1994) asks whether the modern economics that he had so much of a hand in constructing could be effectively used in the service of the nineteenth-century humanitarian goals of socialism. Stiglitz answers in the affirmative, and his work on inequality is a prime example of what he means in making that claim.

At this point, it is important to take a step back and address the exact relationship between ethics and economics that is being articulated. In definitional terms, ethics is the study of the ends for which men should strive, while economics is the study of the relationship between scarce means and unlimited ends. Economics does not discuss ends as such. From these definitions, the relationship between ethics and economics consists of separate evaluations of ends and means (Rothbard 1976). Ethical concerns provide the analyst with the questions to be pursued, while economic analysis provides the answer to the viability or efficacy of means for achieving the lofty goals that ethical discourse sets for individuals (Boettke 2007). This is the relationship between economics and ethics that guided (and continues to guide) the analysis of Austrian school thinkers. Their critique of socialism did not discuss the desirability of socialist ends. Likewise, their discussions of free markets and capitalism concerned the effectiveness of such a system in enabling individuals to achieve their ends without discussing the desirability of those ends.

The existence of a strict separation of economics and ethics has been questioned. Some philosophical critiques of economics as a discipline have suggested that economics is full of ethical content. The very definition of words such as “voluntary” and “coercion,” or even “property rights,” have ethical content embedded in them that cannot be ignored. In discussions of the “just” distribution of property, for example, there are implicit and explicit ethical judgements made throughout that address a range of topics such as acquisition, use, and transference of such property (Rothbard 1973; Block 1975). Economics, by treating the object of study as ethically neutral, reflects an ethical endorsement of the existing pattern of resource ownership (High 1985). A correct counter to this argument, however, is simply to point out that the status quo has no ethical priority. It just is what it is. Thus, the analysis of the situational logic must begin with recognizing the initial starting point.

To quote Pigou again, “If practice is the impulse to the economist’s work, it is plain that Economics cannot stand alone. For our science is not a normative but a positive science” (1908: 13). As he argues, economists cannot tell you what ought to happen, only what is likely to happen given a change in the context of economic decisions. Economics cannot guide reform on its own. It cannot “pass judgements on questions of the relative goodness-in-themselves of various states of conscious life” (ibid.). That is the domain of ethics.

This division of labor would be fine if it were maintained in practice. It is useful only to the extent that analysts understand that the high ideals of the ethicist will only have practical importance to the extent that the “ordinary business of life” is allowed to inform the understanding of the viable as well as the desirable (ibid.). Grand ethical

aspirations that are not economically viable are not simply grand plans that have gone awry. Instead, they are potentially unethical delusions that harm human beings and reduce well-being.

In their development of the strict Weberian idea of economics as *wertfrei*, or value-free, Mises and Robbins directed their criticisms at the theory of welfare economics developed by Pigou. Pigou thought he could develop a science of economics that could provide answers to passionate concerns with the plight of man: we need not rely only on the poets and Romantics to see with “vividness the faces of the suffering and the degraded who have been worsted in the industrial struggle” (1908: 12). Pigou thought he could use Marshallian analysis and his idea of “welfare economics” to design solutions—as a doctor would prescribe a medicine—to cure the ills of industrial society. In his mind, redistribution would bring relief from poverty; regulation and taxation would curb monopoly power and externalities; and taming of industrial fluctuations through judicious intervention would correct for unemployment and labor strife. But precisely *how* is the *science* of economics going to justify the necessary coercive interventions to stave off industrial conflict?

Mises and Robbins sought to address this question by adopting their roles as defenders of the scientific status of economics and as critics of the proposed set of interventions and articulated their case for a value-free economics. Economics as a scientific discipline cannot simply deliver knowledge to those who want it merely because they want it. It has limits to its inquiry just as all scientific disciplines do. It also has more impact in its negative knowledge than most policy activist would like to acknowledge.² Thus, economics puts parameters on utopias. It does this first and foremost by stressing the primordial fact that we live in a world of scarcity. Recognizing the fact of scarcity implies that we are forever negotiating trade-offs, and in that negotiating of trade-offs we must deal with the incentives that actors face in the conduct of the “ordinary business of life” as well as the unique knowledge of circumstances they use in making their choices. High-minded ideals that do not account for the incentives and knowledge involved in the “ordinary business of life” will amount to little more than daydreaming at best, and could inflict harm on those who were the intended target for assistance. None of what we have said so far entails value judgements or ethical assessment. It simply follows from the positive analysis of the efficacy of chosen means to satisfy given ends.

The economist must limit their role to that of a student of society, and, if done with intellectual discipline, a social critic—but a critic that only engages in critique founded in means-ends analysis, and never the transcendent critique of imagined alternative ends to be pursued. Mises and Robbins, following in the utilitarian tradition of English classical political economy, thought that the assumption of universally shared goals for public policy—namely the public interest and in particular the welfare of the least advantaged—resolved many of the thorny issues in utilitarian philosophy. But economists in Cambridge at the time, such as Pigou, thought you could combine that alleged

² By “negative knowledge,” we refer to the point that Hayek (1945) makes, namely that policy planning requires access to knowledge that planners do not and cannot know.

commitment to a universally shared ethic with the public interest using the tools of economic reasoning based on interpersonal comparisons of utility.

Two critical things emerge that threaten the scientific status of economics if Pigovian welfare economics comes to characterize the economic discourse. First, value judgments are smuggled, unexamined, into the analysis at various stages. Second, various unwarranted assumptions and conceptual confusions related to the concepts of utility, equilibrium, and optimality are introduced. The *science* of economics is reduced to normative philosophizing parading as positive analysis. Furthermore, there is absolutely no argument whatsoever that the economist should have any privileged position in deliberations for collective action other than the mantel of science; after that has been abandoned, all that is left is the charisma of the personality.

Mises and Robbins (and Hayek) were embroiled from 1920–40 in various heated public policy debates concerning socialism, war, and depression. They believed that economics had something to offer to these pressing issues, but only if it retained its strict scientific stature. They stressed how alternative institutional arrangements impact the ability of individuals to pursue productive specialization and realize peaceful social cooperation. Clearly defined and strictly enforced private property rights yield one pattern of exchange and production, while poorly defined and weakly enforced property rights yield a different pattern. In theory, collective property rights regimes, *by definition*, would have to forego the economic forces that follow from a regime of property, contract, and consent. Additionally, in practice, a collective property rights regime would deviate from the theory and function of property rights due to attenuated property rights and the disparities of bargaining power that political privilege wrought.

In the debate over market socialism, the first line of argument was the inability for even selfless planners to engage in rational calculation of alternative productive ventures. Mises, Hayek, and Robbins did not emphasize the incentive problems that socialist proposals would face because they wanted to grant, for sake of argument, as much common ground as possible and highlight the decisive objection: socialism, by design, would have to forgo the intellectual division of labor that characterizes a modern economy. Their argument evolved in the 1930s and 1940s into one that stressed the *political consequences* of the planned economy. They sought to demonstrate that socialism was incompatible with democratic aspirations in works such as Hayek's *The Road to Serfdom* (1944) and Mises' *Omnipotent Government* (1944). The British advocates of market socialism often argued that they were socialist in their economics because they were liberals in their politics, so if the argument of Mises and Hayek had force it would be another intellectual death blow to the argument motivating the growth of government planning in the wake of the Great Depression and the Second World War.

This is why Hayek's book is written as a tragic tale: the serfdom that results is the unintended and undesirable consequence of a path that is pursued with the highest of ideals in mind. Mises' story rhetorically is less a tragic tale and more a dire warning. But, both stress how liberal democracy is incompatible with the loss of economic freedom, and attempt to make this argument in a way—read carefully—that is positive and not normative. The argument they present constitutes *is* and *follows from* statements, not

should and *should not* statements, and the notion of *desirable* they employ is from the point of view of the advocate or some widely shared ideal of the public interest. Motives are not questioned. Ends are not subjected to critical dialogue. Instead, the analysis is one of the efficacy of chosen means in the obtainment of given ends.

11.3 ETHICAL CONCERNS, ECONOMIC ANALYSIS

The Austrian argument for the strict value-free nature of economic analysis limits what economics can do. To demand that a discipline do something it cannot is a surefire way to corrupt a discipline. This is why the Austrian writers were adamant that economics cannot be a tool of advocacy. Rather, it is a tool of social understanding and social criticism that provides an analytical understanding of the forces or processes *at work* in any economic system, as well as negative knowledge. Economics cannot answer questions like whether profits are deserved, but its analysis of factor pricing and the allocation of entrepreneurial talent can inform on the consequences of various answers to such questions in relation to the pattern of exchange, production, and distribution.

Ethical concerns can, and do, ignite our imagination as economists such as they do citizens in general. Joseph Schumpeter in his *History of Economic Analysis* (1994), referred to the *pre-analytic cognitive act* that often provides the raw material for scientific analysis as *vision*. Vision, he argued, was a critical part of the scientific-scholarly enterprise. Concerns about poverty, ignorance, mass unemployment, gross injustice, and so forth clearly can ignite—and one could argue always have ignited—the imagination of economic thinkers. There are social ills that must be addressed.

Perhaps the social “miracles” that some economists, such as F.A. Hayek, have identified in the coordination of vast multitudes through market exchange, are merely a precursor to identifying root causes of those social ills. In “The Use of Knowledge in Society” (1945), Hayek described the functioning of the price system as a “marvel” to remind his peers the basic teaching of economics: the unmatched success of the price system in allocating resources to their highest valued use. The basic structure of the economy is the most powerful tool for satisfying individuals’ wants. Additionally, Hayek (1931: 224–5) argued, as a methodological point, that one must understand how a system could in theory achieve full employment before postulating an explanation for unemployment. If a thinker did not follow this procedure, and instead began her analysis with unemployment, then what explanation for unemployment could be offered? None: it would just be the starting point. Unless we can discuss causes of the initial condition, how can we learn more about the social ills that we hope to eradicate, or at least ameliorate.

What is true for the analysis of unemployment, is also true for Want, Disease, Ignorance, Squalor, and Idleness (Beveridge 1942: 6). The economic way of thinking is a powerful tool—perhaps the most powerful tool in the social scientist’s toolkit. It enables

a properly trained thinker to assume the observational heights of a genius through its use, while a genius working without this tool mistakes nonsense for analysis (Buchanan 1966: 7). As the intellectual stakes rise—like in studies of development—the consequences of this mistake in thought rise exponentially. Human suffering and misery, we contend, is largely man-made. No doubt there are geographic advantages in development, as there are random events of purely natural origins that can bring in their wake destruction and devastation, but the vast majority of human suffering is due to the cruelty of man, and driven more often than not by bad ideas that work in concert with special interests.

F.A. Hayek argued in his Nobel Prize lecture (1974), and elsewhere, that nobody is as dangerous as an economist who was only an economist. Hayek was making an appeal for economic science to be embedded in politics and law, as well as culture and history. It was not a denial of the universal applicability of the logic of choice, or what Hayek called “the economic calculus.” Rather, it was a recognition of the context specific nature of any efforts in applied economics, as well as a corresponding call for humility among economists with respect to their ability to manipulate that context. Hayek concluded his Nobel address by arguing that if the discipline of economics is to do more good than harm, economists must cease envisioning their task as engineers fine-tuning a machine. Instead, they must focus their attention on understanding what alternative institutional arrangements cultivate and facilitate the creative powers of a free civilization with its corresponding growth of knowledge and, thus, of wealth.

Similarly, James Buchanan (1987) argued that economists should stop giving advice as if they were offering it to a benevolent despot. In his effort to balance the economics discipline’s attraction to the zealot and reformer with the demands of objectivity and rigorous analysis, Buchanan (1959) introduced two stages of analysis and an insistence on the priority of democracy. In his system there is the analysis of the rules of the game, and then an analysis of the best strategies players will play given that set of rules. What we conjecture to be a “good game” results from a critical analysis that tacks back and forth between these two levels of analysis.

Based on this exercise, Buchanan argued, we could, as citizen-economists who have no claim whatsoever to any status above our fellow citizens, offer the suggested structural rule changes that our analysis leads us to believe will be Pareto improving, to “test” in the democratic process of collective deliberations (1959: 137).

Propositions advanced by political economists must always be considered as tentative hypotheses offered as solutions to social problems. The subjective bases for these propositions should emphasize the necessity for their being considered as alternatives which may or may not be accepted. But this is not to suggest that one proposition is equally good with all others... But this change is a ‘cure’ [to social ills] only if *consensus* is attained in its support. (ibid.: 128)

Let’s make this as clear as possible: economists are not making recommendations for in-rule changes in economic activity. All suggestions for changes are at the level of

structural rules. But those changes can never be imposed as part of some expert demand to be undeterred by democratic checks. Instead, rule changes are “tested” in the democratic process of collective decision making to see if we can secure agreement from fellow citizens. This securing of agreement means that the threat of externalities in the political process will have been minimized.

If we hope to enlist our vision of social ills to direct, and hopefully improve, our positive economic analysis, then either we go in the direction of Hayek and Buchanan in terms of a focus on rules, or we have to postulate some universal ethic that is widely shared as some ultimate given end (Rothbard 1976). If we admit that there exists a universal ethic—either derived from rational analysis such as in the natural law tradition (for example, Rothbard and the non-aggression axiom) or simply postulated as an appeal to common-sense (such as Kirzner and the finder’s keeper ethic)—then we use positive economic analysis to inform us about the impact of alternative institutional arrangements on our ability to realize these ethical ends.

None of the approaches avoid tensions at a conceptual and practical level, but tensions probably plague all efforts at theorizing about the human condition and the ethical assessment of the just and the good. While Hayek may have warned about the danger of the economist who only knows economics, surely the most dangerous of all intellectual creatures is the moral philosopher who knows no economics. Economics puts parameters on our utopias. As economists, we do not only study choice within constraints; the discipline enables us to understand constraints on our mental flights of moral fancy. We must strive for a morality for mortals, not angels.

In one of the smartest essays by an economist on this subject, Dennis Robertson (1956) famously asked “What Does the Economist Economize?” His answer was that we economize the most precious of all resources—love. If our philosophical ideals require us to postulate that mankind must love one another for them to work, they will exhaust the supply of love and prove them to be unobtainable, resulting in human suffering and misery. Instead, we must strive to find institutional arrangements that improve the human condition and do not rely on love for their operation, but instead invoke ordinary human motivations and cognitive capacities. In the world of mortals, omniscience, omnipotence, and benevolence are not acceptable assumptions. Invoking them may produce romantic poetics, but not political economy or social philosophy.

But there should be no doubt that economists—as citizens and as scientists—can, and in fact often will, have serious ethical concerns that motivate their scientific inquiry. That is, we contend, not only a natural human impulse, but a healthy aspect about the “worldly philosophy.” The effort to eliminate ethical concerns from economic discourse was a huge intellectual error, which did not result in ethical concerns disappearing from economics, but in these ethical concerns being smuggled into the analysis in extremely unhealthy and unacknowledged ways. Amartya Sen’s *On Ethics and Economics* (1987) argued persuasively for a useful way to think about production of economic thought as in-line with a production possibility frontier, with economics as engineering on one axis and economics as ethics or moral philosophy on the other. Clearly, Sen suggests, the economics profession moved toward a corner solution of economics as engineering

throughout the twentieth century. He argues it is time to perhaps move along the frontier and away from the corner solution by opening up professional discourse among economists of an ethical or moral philosophical nature. Subsequent works by scholars such as Deirdre McCloskey, as well as Sen himself, suggest that such a movement along the frontier need not produce crippling trade-offs, but perhaps aid in the process of pushing the frontier of economic knowledge production farther out from the origin (McCloskey 2010). In short, the *science* of economics improves with a return to the *moral philosophy* questions of Hume and Smith, of Say and Bastiat, of Mill and Wicksteed, of Mises and Hayek, of Buchanan and Sen, of McCloskey and Schmidt, Gaus, Storr, and Chamlee-Wright.

11.4 SOCIAL COOPERATION IN THE AUSTRIAN SCHOOL

The Austrian school engages those moral philosophy questions while maintaining a *wertfrei* analysis through an emphasis on social cooperation. For Mises, social cooperation is the only ethical standard by which social rules and policies can be judged: “There is no other standard of what is morally good and morally bad than the effects produced by conduct upon social cooperation” (1952: 145). By holding maximizing social cooperation as the end for which economists analyze the purposed means, economists can pursue an answer to why some nations are rich and others are poor without sacrificing individuals’ ends to those of “society”—or more accurately, those of the economist. An emphasis on social cooperation enables the economist to remain rigidly value-free, and focus her analysis on the effect of policies on social cooperation in a society.

From Philip Wicksteed in 1910 to Tyler Cowen today, Austrian school thinkers have emphasized how social cooperation is the path to prosperity. Prosperity depends on structural rules that enable individuals to cooperate in the achievement of their diverse ends. For this reason, Austrian school thinkers emphasize an institutional framework of property, contract, and consent—all of which constitute a market system.³ Such structural rules enable individuals to exchange freely. Through exchange they aid in the achievement of each other’s ends. Thus, an important aspect of the market system is that it is means-based, not ends-based, meaning that individuals need not have the same goal. Instead, individuals use the price system and exchange in order to achieve their ends while meeting the ends of others. This process of mutual satisfaction through social cooperation aims at the long-run maximization of individuals’ ends, and, ultimately, economic growth.

³ Following Kirzner (1989: 8), we understand the “market system” as identical with “capitalism” and describing a society characterized by specific structural rules—namely private property rights, freedom of contract and consent, and voluntary participation.

In *The Common Sense of Political Economy* (1910), Wicksteed introduces *non-tuism* as a central aspect of the market system. An individual need not condone the ends of others or their moral character to satisfy their ends while pursuing her own. The vicious and virtuous serve each other. Nor does an individual need to be self-sufficient. Cooperation through the market system allows her to achieve her ends without sharing the values or resources of the people with whom she cooperates.

And so, under the all-covering cloak of money payments for services and commodities, and sales of instruments and supports of life for money payments, all purposes and impulses, of love and of lust, of narrow greed and of broad beneficence, of enlightened and productive insight, of blind, tangled, and self-confuting gropings, all destructive and reckless passions, all wasteful and desolating vices, all noble ambitions, all vulgar or refined enjoyments, all fruitful enterprises, and all foolish or wicked schemes of industrial waste, enter the open market and draw to themselves the efforts and services of men in proportion not to their worthiness or fruitfulness, but to the means they command of furthering the purposes of others; for they secure the co-operation of all sorts and conditions of men, not in the measure in which such men sympathise with them, but in the measure in which by serving them they will forward their own purposes. Neither the urgency of his wants nor the nobility of his purpose determines the extent to which a man may rely on economic forces to help him. (1910: 395–6)

Cooperation through the division of labor and the market process leads to greater prosperity, meaning that more ends are met at the sacrifice of fewer means. The results of the market process do not depend on one's being judged as morally good. Rather, satisfaction of ends relies on one's ability to satisfy the ends of others.

Mises builds on the non-tuism of the market process that Wicksteed highlighted. He notes that individuals do not serve each other in the market because of feelings of benevolence. Instead, Mises posits that feelings of benevolence develop due to the operation of the market system, and its ability to bring about the achievement of diverse and occasionally competing ends.

Within the frame of social cooperation there can emerge between members of society feelings of sympathy and friendship and a sense of belonging together... However, they are not, as some have asserted, the agents that have brought about social relationships. They are fruits of social cooperation, they thrive only within its frame; they did not precede the establishment of social relations and are not seed from which they spring. (1949: 144)

For Mises, like Wicksteed, non-tuism characterizes economic relationships. It is this character that allows for social cooperation and greater prosperity.

Uneasiness, or a longing for greater prosperity, is what initially drives individuals to seek cooperation. Cooperation through the division of labor arises to satisfy individuals' felt unease, and, thus, allows individuals to achieve higher productivity, and enjoy greater prosperity.

In order to comprehend why man did not remain solitary, searching like the animals for food and shelter for himself only and at most also for his consort and his helpless infants, we do not need to have recourse to a miraculous interference of the Deity or to the empty hypostasis of an innate urge toward association. Neither are we forced to assume that the isolated individuals or primitive hordes one day pledge themselves by contract to establish social bonds. The factor that brought about primitive society and daily works toward its progressive intensification is human action that is animated by the insight into the higher productivity of labor achieved under the division of labor. (Mises 1949: 159)

Mises highlights that no contract was necessary to achieve a society based on cooperation through the division of labor. Rather, the unease—the desire for greater satisfaction of wants—that underlies human action motivated individuals to cooperate and develop a society based on a division of labor.

Hazlitt builds on Mises' conception of the morality that arises from the market system. Like Mises, he emphasizes that cooperation is the means by which individuals achieve their ends. This cooperation requires no unanimity in ends or values. It is independent of one's moral character or goal. Hazlitt extends Mises' analysis by discussing how the diverse and competing ends of individuals are not in conflict with society as is often argued. Instead, society and the individual are in harmony.

The aim of each of us is to maximize his own satisfaction; and each of us recognizes that his satisfaction can best be maximized by cooperating with others and having others cooperate with him. Society itself, therefore may be defined as nothing else but the combination of individuals for cooperative effort. If we keep this in mind, there is no harm in saying that, as it is the aim of each of us to maximize his satisfactions, so it is the aim of 'society' to maximize the satisfactions each of its members, or, where this cannot be completely done, to try to reconcile and harmonize as many desires as possible, and to minimize the dissatisfactions or maximize the satisfactions of as many persons as possible in the long run. (Hazlitt 1999: 10)

Hazlitt further extends Mises' analysis by engaging directly with ethics, which "is not a science of what is or was but of what ought to be" (1999: 7). However, embracing value-freedom in economics requires him, and the Austrian school, to accept individuals' desired ends as what ought to be. Thus, he says, because social cooperation through the division of labor is what allows most individuals to achieve their desired ends, the market system must be at the core of "a rational system of ethics" (*ibid.*: 9). Choosing any other system of rules would involve choosing which ends are better for individuals to pursue. Through interaction in the market process, individuals (as society) can choose which ends best forward themselves and society, since only those who serve others in pursuit of their ends achieve those ends. Emphasis on social cooperation is the heart of the connection between economics and ethics in the Austrian school. Holding increased cooperation as an end is what enables more people to achieve more of their desired ends, and thereby enables the greatest prosperity in the long run.

For each of us social cooperation is the great means of attaining nearly all our ends. For each of us social cooperation is of course not the ultimate end but a means. It has the great advantage that no unanimity with regard to value judgements is required to make it work. But it is a means so central, so universal, so indispensable to the realization of practically all our other ends, that there is little harm in regarding it as an end-in-itself, and even in treating it as if it were the goal of ethics. In fact, precisely because none of us knows exactly what would give most satisfaction or happiness to others, the best test of our actions or rules of action is the extent to which they promote a social cooperation that best enables each of us to pursue his own ends. (1999: 12)

Hayek, like Wicksteed, Mises, and Hazlitt, emphasizes how the market system enables individuals to satisfy their often-conflicting ends by serving the ends of others. Having different ends is what ultimately enables individuals to satisfy those of others. If everyone had the same ends, exchange would be impossible in a world of scarcity. Thus, Hayek argues, prosperity is only possible because the market system is means-based, not ends-based; non-tuism is central to the operation and success of the market system.

The decisive step which made such peaceful collaboration possible in the absence of concrete common purposes was the adoption of barter or exchange. It was the simple recognition that different persons had different uses for the same things, and that often each of two individuals would benefit if he obtained something the other had, in return for his giving the other what he needed. All that was required to bring this about was that rules be recognized which determined what belonged to each, and how such property could be transferred by consent. There was no need for the parties to agree on the purposes of each partner in the transaction, and that they thus assist the parties as means for different ends. The parties are in fact the more likely to benefit from exchange the more their needs differ. While within an organization the several members will assist each other to the extent that they are made to aim at the same purposes, in a catallaxy they are induced to contribute to the needs of others without caring or even knowing about them. (1977: 109)

The cooperation and productivity of individuals does not depend on having the same ends, but rather using the same means—exchange—to achieve those ends. Crucially, such exchange must be couched in the right structural rules: property, contract, and consent.

In contrast to other Austrian thinkers who emphasize the role of the division of labor in facilitating cooperation, Hayek emphasizes the ability of the market process to bring about cooperation by reconciling diverse knowledge and plans.

The important point about the catallaxy is that it reconciles different knowledge and different purposes which, whether the individuals be selfish or not, will greatly differ from one person to another. It is because in the catallaxy men, while following their own interests, whether wholly egotistical or highly altruistic, will further the aims of many others, most of whom they will never know, that it is as an overall

order so superior to any deliberate organization: in the Great Society the different members benefit from each other's efforts not only in spite of but often even because of their several aims being different. (1977: 110)

This difference in knowledge and plans is what allows individuals to satisfy the ends of others. Thus, it is crucial for long-run prosperity that society be means-based rather than ends-based.

Building on a tradition of Austrian emphasis on social cooperation through the market process, Cowen (2017) argues that our concern, as economists and citizens, should be maximizing long-run growth. It is through "sustainable economic growth" that we are happier, healthier, and wealthier. A focus on the long run is a key element of the Austrian emphasis on social cooperation. Hazlitt anticipates Cowen when he writes:

Ethics must take into consideration not merely the immediate but the longer effects of any act or rule of action; it must consider the consequences of that act or rule of action not merely for the agent or any particular group but for everybody likely to be affected, presently or in the future by that act or rule of action. (1999: 31n1)

Calling himself "two-thirds a utilitarian," Cowen argues that we need to maximize growth according to an ethical framework based on rules and rights. Achieving Cowen's goal of maximized long-run growth requires that we continue to rely on the rules and rights of the market system and the cooperation inherent to it. A system that maximizes social cooperation will achieve the long-run economic growth for which Cowen (2017) calls.

11.5 CONCLUSION

Economics and ethics study different parts of the same puzzle: human prosperity. Ethics sets the ends for which people should strive, while economics evaluates the viability of the means chosen to achieve those ends. The Austrian school epitomizes this distinction. Austrian school thinkers explore the marvel of human cooperation under the division of labor, and coordination according to the price system. These thinkers maintain their objectivity by adopting a stance of radical subjectivity. They take individuals' evaluation of their own ends as the only relevant valuation. It is not for the economist to declare which ends should be pursued. Thus, the Austrian thinkers' emphasis on social cooperation maintains the individual and her values and ends, as well as encourages a system of rules that leads to long-run growth. In this way, it eliminates any perceived conflict between society and the individual without the economist asserting her values onto individuals or society.

It was through this emphasis on social cooperation that Mises, Hayek, and Robbins maintained their roles as value-free scientists while being perceived as advocates of free markets and capitalism. Rather than submitting their analysis to ideological priors—as is often charged—they embraced objectivity through radical subjectivism.

It is the role of the individual and society, they argued, to determine desired ends through interaction in the market process. Thus, they acknowledged their limited role as students of civilization, leaving the decision of desirability of ends to the ethicist and citizen.

Recognition of the limited applicability of economists' tools—and the limited role economists can hold—requires a re-evaluation of the practice of economics today. We must not ask economics to “bear fruit.” Economics can inform the method of achieving the ends that ethics sets for society, as well as the viability of those ends, but it cannot determine the desirability of those ends. That is the role of the ethicist.

However, the work of the political economist can have value implications. Getting at those value implications requires a Philosophy, Politics, and Economics (PPE) approach (Boettke 1998: 213). Economics and philosophy need to be reunited in the vein of Adam Smith, such that the economist becomes a moral philosopher. This is not to confuse the use of economic and ethical analysis. Each tool must be used within its limits, and the use of each must be explicitly acknowledged. Disciplined practice of economics in a PPE framework will bring unity to the right and the good, encouraging a virtuous and prosperous society.

REFERENCES

- Bentham, Jeremy. 1843. “Anarchicall Fallacies; Being an Examination of the Declarations of Rights Issued During the French Revolution.” In John Bowring (ed.), *The Works of Jeremy Bentham*, Vol. 2: *Principles of Judicial Procedure* (Edinburgh: William Tait, 2011), pp. 896–953.
- Beveridge, Sir William. 1942. *Social Insurance and Allied Services*. London: His Majesty's Stationary Office.
- Block, Walter. 1975. “Value Freedom in Economics.” *The American Economist* 19: 38–41.
- Boettke, Peter Joseph. 1998. “Controversy: Is Economics a Moral Science? A Response to Ricardo F. Crespo.” *Journal of Markets and Morality* 1: 212–19.
- Boettke, Peter Joseph. 2007. “From Approximate Value Neutrality to Real Value Relevance: Economics, Political Economy, and the Moral Ecology of the Market Order.” *Faith and Economics* 49: 19–34.
- Buchanan, James McGill. 1959. “Positive Economics, Welfare Economics, and Political Economy.” *Journal of Law and Economics* 2: 124–38.
- Buchanan, James McGill. 1966. “Economics and Its Scientific Neighbors.” In *The Collected Works of James M. Buchanan*, Vol. 17: *Moral Science and Moral Order* (Indianapolis, IN: Liberty Fund, 2001), pp. 3–23.
- Buchanan, James McGill. 1987. “The Constitution of Economic Policy.” *Science* 236: 1433–6.
- Chang, Ha-Joon. 2014. *Economics: A User's Guide*. New York: Bloomsbury Press.
- Colander, David C. 2000. *Economics*. New York: McGraw-Hill College.
- Cowen, Tyler. 2017. *Stubborn Attachments: A Vision for a Society of Free, Prosperous, and Responsible Individuals*. Medium.com.
- Dekker, Erwin. 2016. *The Viennese Students of Civilization: The Meaning and Context of Austrian Economics Reconsidered*. New York: Cambridge University Press.
- Gunning, James Patrick. 2005. “How to Be a Value-Free Advocate of Laissez Faire: Ludwig von Mises's Solution.” *American Journal of Economics and Sociology* 64: 901–18.
- Hayek, Friedrich August von. 1931. *Prices and Production*. London: Routledge (2008 edition).

- Hayek, Friedrich August von. 1944. *The Road to Serfdom: Text and Documents—The Definitive Edition*. In Bruce Caldwell (ed.), *The Collected Works of F.A. Hayek*, Vol. 2. Chicago, IL: University of Chicago Press, 2007.
- Hayek, Friedrich August von. 1945. "The Use of Knowledge in Society." *American Economic Review* 35: 519–30.
- Hayek, Friedrich August von. 1974. "The Pretence of Knowledge." *American Economic Review* 79 (6): 3–7.
- Hayek, Friedrich August von. 1977. "The Market Order or Catallaxy." In *Law, Legislation, and Liberty*, Vol. 2: *The Mirage of Social Justice* (Chicago, IL: University of Chicago Press, 1978), pp. 107–32.
- Hazlitt, Henry. 1999. *Rules for Living: The Ethics of Social Cooperation, An Abridgement of The Foundations of Morality*. Bettina Bien Greaves (ed.). Irvington-on-Hudson: Foundation for Economic Education.
- High, Jack. 1985. "Is Economics Independent of Ethics?" *Reason Papers* 10: 3–16.
- Kirzner, Israel Meir. 1989. *Discovery, Capitalism, and Distributive Justice*. New York: Basil Blackwell (2016 edition).
- McCloskey, Deirdre Nansen. 2010. *Bourgeois Dignity: Why Economics Can't Explain the Modern World*. Chicago, IL: University of Chicago Press.
- Mises, Ludwig von. 1920. "Economic Calculation in the Socialist Commonwealth." In Friedrich August von Hayek (ed.), *Collectivist Economic Planning* (New Jersey: Kelley Publishing, 1975), pp. 87–130.
- Mises, Ludwig von. 1944. *Omnipotent Government: The Rise of the Total State and Total War*. Indianapolis: Liberty Fund (2012 edition).
- Mises, Ludwig von. 1949. *Human Action: A Treatise on Economics*. New Haven, CT: Yale University Press (1998 Edition).
- Mises, Ludwig von. 1952. *Planning for Freedom and Twelve Other Essays and Addresses*. South Holland, IL: Libertarian Press.
- Myrdal, Gunnar. 1929. *The Political Element in the Development of Economic Theory*. Cambridge, MA: Harvard University Press (1954 edition).
- Pigou, Arthur Cecil. 1908. *Economic Science in Relation to Practice*. London: Macmillan and Co.
- Robbins, Lionel. 1932. *An Essay on the Nature and Significance of Economic Science*. New York: New York University Press (1984 edition).
- Robertson, Sir Dennis Herbert. 1956. "What Does the Economist Economize?" In *Economic Commentaries* (London: Staples Press Limited), pp. 147–54.
- Rothbard, Murray Newton. 1973. "Value Implications of Economic Theory." *The American Economist* 17 (1): 35–9.
- Rothbard, Murray Newton. 1976. "Praxeology, Value Judgments, and Public Policy." In Edwin Dolan (ed.), *The Foundations of Modern Austrian Economics* (Kansas City, KS: Sheed and War), pp. 89–111.
- Schumpeter, Joseph Alois. 1994. *History of Economic Analysis*. Abingdon, UK: Routledge.
- Sen, Amartya Kumar. 1987. *On Ethics and Economics*. Oxford and New York: Basil Blackwell.
- Smith, Adam. 1759. *The Theory of Moral Sentiments*. Dugald Stewart (ed.). London: Henry G. Bohn (1853 edition).
- Smith, Adam. 1776. *An Inquiry into the Nature and Causes of the Wealth of Nations*. Edwin Cannan (ed.). London: Methuen & Co., Ltd (1904 edition).

- Stewart, Dugald. 1793. "Account of the Life and Writings of Adam Smith LL.D." In *The Collected Works of Dugald Stewart, Vol. 10* (Edinburgh: T. Constable and Co., 1854), pp. 1–98.
- Stiglitz, Joseph Eugene. 1994. *Whither Socialism?* Cambridge, MA: MIT Press.
- Weber, Max. 1917. "The Meaning of 'Ethical Neutrality' in Sociology and Economics." In Edward A. Shils and Henry A. Finch (eds.), *The Methodology of the Social Sciences* (Glencoe, IL: The Free Press, 1949), pp. 1–47.
- Wicksteed, Philip. 1910. *The Common Sense of Political Economy*. London: Macmillan and Co. (2003 edition).

CHAPTER 12

FEMINIST ECONOMICS AND ETHICS

ULRIKE KNOBLOCH

BASED on different normative foundations, a plurality of approaches to feminist economics has been developed since the 1980s. The major tasks of an ethics of feminist economics, or *feminist economic ethics*, are to make these normative foundations visible and to critically reflect them from a non-androcentric moral point of view that has first to be unfolded. In doing so, feminist economic ethics integrates the three disciplines of economics, ethics, and feminist theory.

12.1 FEMINIST ETHICS

Feminist ethics looks at contemporary ethical approaches from a gender perspective (Section 12.1.1), makes visible the often implicit gender norms underlying these approaches (Section 12.1.2), and focuses on non-androcentric approaches to ethics, such as an integrative ethics of justice and care (Section 12.1.3).

12.1.1 Beyond Androcentric Ethics

Male-centered or *androcentric* ethics takes the masculine as the criterion of valuation.¹ Being aware that thinking and acting is always situated, feminist ethics starts from the differences in the socially constructed lives of men and women. In overcoming androcentrism, feminist ethics is not relativist, but looking for universal criteria of evaluation and cross-cultural legitimation of norms. On this basis, traditional ethical approaches can be critically reflected and developed further from a feminist perspective.

¹ The term 'androcentric' actually dates back to the early twentieth century (Gilman 1911).

In the following, first the development of the moral point of view and second the women's rights debate, are taken as examples of how to overcome androcentric ethical assumptions.

In developing the moral point of view, the Swiss economic ethicist Peter Ulrich (2008: 43–78) distinguishes several steps with regard to the evolvement of reciprocity. The first step is the commonly known *golden rule* of reciprocity: “Treat others as you would like to be treated.” Yet in this case, reciprocity goes only in one direction: an action is judged according to whether oneself would like it, leaving out the opposite direction. Adam Smith's *impartial spectator* includes the perspective of another yet abstract person: “We endeavour to examine our own conduct as we imagine any other fair and impartial spectator would examine it” (1759: 110). Immanuel Kant refines Smith's moral principle into the *categorical imperative*: “So act that you use humanity, whether in your own person or in the person of any other, always at the same time as an end, never merely as a means” (1785: 80). The last step in developing the moral point of view, according to Ulrich, is the regulative idea of the ideal communication community of discourse ethics by Jürgen Habermas and Karl-Otto Apel.

The development towards a feminist-critical moral point of view, however, does not stop here. Feminist discourse ethics seeks to overcome the implicit androcentrism of abstract discourse principles. Taking a closer look at these principles, especially the situation of the discourse participants, Seyla Benhabib has distinguished a generalized and a concrete other as two necessary perspectives and formulates a “moral standpoint of care” (1992: 185) as the moral standpoint of connected subjects, rather than separative selves.

Starting from human lives, the idea of human rights is meant to protect and support the lives of *all* human beings. To overcome “the androcentric structure which shaped the development and expression of human rights from its very beginning” (Heimbach-Steins 2009: 251), their formulation has to reflect on the concrete lives of women as well. Women's rights are focused on the danger to life of women as women concerning all three groups of human rights: personal, political, and ESC (economic, social, and cultural) rights. It is still a challenging task to increase the effectiveness of the United Nations covenants with their monitoring processes; related to feminist economics, these are especially CEDAW (Convention of the Elimination of all Forms of Discrimination against Women, 1979) and ICESCR (The International Covenant on Economic, Social and Cultural Rights, 1966).² Furthermore, societal as well as philosophical discussions about not yet codified, and still unknown human rights are required, especially about care rights and the rights of people with dementia and Alzheimer's disease.

12.1.2 Critical Reflection of Gender Norms

Norms provide orientation about the right and the good for an individual or a group. In turn, norms—including gender norms—are socially constructed and deeply rooted in the

² Despite its great significance, the ICESCR is rightly criticized because it does not “recognize the needs of the many women who do not participate in the paid economy, or participate as unpaid family labor in family businesses and who are thus not covered by the rights that are conferred on workers who earn wages in the paid economy” (Elson and Gideon 2005: 23).

institutions of every society and economy. Gender norms mean “differential rules of conduct for women and men, including rules governing interactions *between* women and men” (Pearse and Connell 2016: 35) and include all privileges related to biological sex, such as the *patriarchal dividend* (as coined by Connell), capturing for example that men do less unpaid work than is done for them. Like all other norms, gender norms are ambivalent: they can either hinder or support gender equality (Pearse and Connell 2016: 35) and sometimes they function “both ways at once” (Butler 2004: 8). Also, economic thought and action are based on gender norms usually neither made visible nor critically reflected upon.

To be able to explain gender inequalities in economic processes and their persistence, a critical reflection of the normatively structured economy and economic system is needed, addressing three levels: at the *individual* level, gender norms are scrutinized; at the *social* (later called *organizational*) level, gender relations, especially the gender-based division of labor, are examined; and at the *structural* level, the gender regime, understood as the institutionalization of gender norms and gender relations in the societal structure from marriage laws to social security, is explored.

12.1.3 Care Justice for Whom?

Care ethics is the most prominent approach to feminist ethics and has influenced many disciplines and societal discourses. But feminist ethics as such started much earlier with the existential ethics in Simone de Beauvoir’s *The Second Sex* (1949), and represents a much broader debate, including Luce Irigaray’s *The Ethics of Sexual Difference* (1984). However, it was Carol Gilligan’s empirically based book *In a Different Voice* (1982) that launched the discussion about the existence of one universal moral versus two distinct morals, a male ethics of justice and a female ethics of care. After years of heated debate, this question of moral development seems to have been more or less resolved by conceding that every human being is able to develop into not only one, but (at least) two directions, a moral development in the direction of justice *as well as* in the direction of care.

The next generation of care ethicists have discussed care not only at the personal and national, but also at the political and global level. Joan Tronto (1993, 2013) and Selma Sevenhuijsen (1998) have developed a *political* theory of care that rethinks “liberal concepts such as equality, justice and autonomy . . . from the perspective of the ethics of care” (ibid.: 34). Fiona Robinson (1999, 2013) has outlined a *global* theory of care scrutinising inequalities in doing care worldwide. To her,

care provides a concrete focus for thinking about injustice on a global scale by revealing the (complex) reasons why people lack adequate care, why they are unable to give and receive care effectively and why caring labor is often unrecognized, exploitative and undertaken at great cost by the most marginalized members of societies. (2013: 132)

Today the approaches to care ethics include other important questions, such as caring rationality (Wærness 1984), the right to care or not to care (Knijn and Kremer 1997), caring

citizenship (Sevenhuijsen 1998), caring masculinities (Hanlon 2012), caring democracy (Tronto 2013), and the moral harm of care migration (Kittay 2014).

Concerning the necessary *integration* of the ethics of justice and the ethics of care, Christa Schnabl (2005: 247–94) identifies and systematizes three types of approaches. For the *dualistic* approach, care ethics has primacy over the ethics of justice, such as in Nel Noddings (1986). The *corrective-integrative* approach includes care into the primary ethics of justice, such as in Susan Moller Okin (1989). Finally, the *transformative-alternative* approach outlines a justice-based care ethics starting from the practical experience of care and its special characteristics, such as in Eva Feder Kittay (1999).

The *ethical turn* in feminist theory (Garber et al. 2000) extends the reflexivity of feminist ethics in at least two additional directions. Considering the still omnipresent two-gender system, a *queer ethics* uses the term *heteronormativity* (Warner 1991) as critique of this system. Heterosexual normativity structures everyday life and society, making heterosexuality with its institutionalized manifestation a norm for everyone. Referring to Judith Butler (2004, 2005), Annika Thiem outlines Butler's queer ethics as “performed through her mode of inquiry, which is committed to beginning from the margins and the marginalized to query the social norms and structures that condition, enable, and animate forms of marginalization” (2008: 8). The other necessary extension of reflexivity relates to *feminist postcolonial ethics* and its “anti-essentialist ethical universalism . . . doing more than critiquing the West, offering it a different way of thinking about the sources and commitments on which ethics can rest” (Charusheela 2004: 56; see also Mohanty 2003).

Even after including queer and postcolonial ethics, other perspectives, categories and dimensions of reflection not yet covered are to be expected and, over time, have to be conceptually included into a critical reflexive feminist ethics. Searching for such a cross-cultural feminist ethics, aware of missing out cultural and other influences, is an everlasting search process.³ An ethics of caring provisioning could be a next step (see Section 12.3.3).

12.2 UNDOING GENDER IN ECONOMICS

12.2.1 Degendering Economic Terms

In this section, the normative content of some basic economic terms is carved out from a gender perspective: first, economics and the economy; and second, work and labor.

The term *economics*, as the study of the *economy*, can be traced back to the ancient Greek term *oikonomia*, consisting of *oikos*, meaning “house” or “housing community,”

³ The Swiss social ethicist Arthur Rich (1991: 19, own translation) wrote: “A conscientious ethics can never be presented as a finished and final value system that is able to offer ready-made prescriptions for actions. It is and will always remain a task that never comes to an end, a demand that is never pacified with a definitive answer.”

and *nemein*, meaning “to allot” or “to organize.” *Oikonomia*, accordingly, was the art of managing the household, whereas dealing with more money than needed in the household was considered *chrematistics*. At the end of the sixteenth century, a second understanding of the term “economy” evolved: *political economy*. James Steuart wrote in *An Inquiry into the Principles of Political Economy* (1767: 2): “What œconomy is in a family, political œconomy is in a state.” Even if he still used economy in the first sense in *The Wealth of Nations* (1776), Adam Smith was focused on political economy and the newly appearing market economy, not on household management.⁴ Yet, a market economy is always “embedded” (Polanyi 1944) into “the other economy” (Donath 2000), based on power structures and regulated by the state. Feminist economics, therefore, has to include a critical reflection of the status quo as well as the power and gender relations a market theory is based upon (see Anderson 1993).

Concerning human activities, many European languages distinguish between two terms: *work* as creative and fulfilling doing, and *labor* as hard and physical activities. Over the past centuries these terms have changed their meaning significantly (Komlosy 2014). Only since the end of the nineteenth century, they are both understood as income-generating activities with strong societal consequences: “The definition of work as an employment outside the home, paid, legally codified, institutionalized and socially secured determines the whole set of rules of the industrial society” (ibid.: 11–12, own translation). Feminist economics is based on the assumption that all forms of unpaid work and labor have to be included into economic analysis because they are necessary for every society to survive and live well together. In time-use studies carried out by the national statistical offices in many countries since the 1990s, unpaid activities are distinguished into housework, care work and voluntary work and defined with the help of the *third-person criterion* (Reid 1934: 11), according to which unpaid work includes all activities that could be done by a person outside the own household. These studies have shown that, on the national basis, unpaid work takes at least the same amount of time as paid work, and that women do most of the unpaid work whereas men do most of the paid work. Moreover, the social distribution of labor does not only go along lines of gender, but also along categories of class, race, nationality, age, and other categories of intersectionality.⁵

12.2.2 Normative Foundations of Feminist Economics

All approaches to feminist economics include gender as a category, recognize unpaid work as work, and analyze what it means that it was assigned to women. But the often implicit normative assumptions are different. Out of the plurality of feminist economic

⁴ According to the etymology of the term *economy*, the term *household economy* is a pleonasm.

⁵ There are many other economic terms that can be examined for their gender normativity, including rationality, productivity, efficiency, value, utility, preferences, profit, production, reproduction, and consumption; space constraints prohibit their discussion here.

approaches, assumptions will be identified for the three most common approaches: feminist neoclassical, feminist Marxist, and feminist institutionalist economics (see Hoppe 2002).

Feminist neoclassical economists stick with the assumptions of neoclassical economics and their normative foundation to address issues like household production and the gender-based division of housework.⁶ However, some of these assumptions have an androcentric bias (see Ferber and Nelson 1993, 2003; Kuiper and Sap 1995): *homo economicus* is the construction of an independent, unembedded economic agent without unpaid care responsibilities. *Methodological individualism* looks at society from the perspective of independent individuals choosing between market options. The underlying theory of the good is based on *utility* interpreted not as happiness, pleasure, or desire fulfillment, but as the fulfillment of given preferences via the market as good in itself. Neoclassically based *neoliberal economic policy* believes that free markets lead to the best results for all people and does not critically reflect on the gendered structure of the status quo, such as the gender distribution of income and property rights.

Feminist Marxist economists use historical materialism to explain the oppression and exploitation of women in the capitalist economy. The Marxist-based *domestic labor debates* claim to systematically include housework into the discussion on the process of accumulation and its structures of power. According to Marx's labor theory of value, a "use value, or useful article... has value only because human labor in the abstract has been embodied or materialized in it" (1867: 53). One of the big questions in this context was whether and how housework adds and should add to the creation of surplus value. In her *dual systems theory*, Heidi Hartmann (1979: 2) has shown that modern society "is organized both in capitalistic and in patriarchal ways," with these systems operating sometimes in tandem and other times in opposition. To look ahead to the Marxist vision of a classless society from a feminist perspective, it should be kept in mind that "what really matters for the economy is the way norms are materialized in social life" (Pearse and Connell 2016: 37), and that the material turn to a new feminist materialism brings new challenges beyond the critique of power relations and their material structure (see Bauhardt 2017).

Feminist institutional economists are even more diverse, ranging from the works by Ann Jennings (1993) to Zdravka Todorova (2014). They often start with Thorstein Veblen (1899, 1904, 1914) or his successors, but the German Historical School could also be taken as a starting point. According to the institutional perspective, the life of every individual is embedded in institutions which are understood as societal constructions that shape the acting and agency of human beings. At the same time, these institutions are made by human beings and, therefore, can be changed by them. So, institutions support and constrain human activity and agency at the same time. Feminist institutional economics helps to degender the processes of valuation in the economy, to make visible the institutionalized gender norms, relations, and regimes (Mayhew 1999: 737). If the value of work

⁶ New Institutional Economics (NIE) and New Home Economics (NHE) belong to neoclassical economics as well.

is connected to income, unpaid work is not adequately valued and, thereby, the value of housework is not visible in the economy.

12.2.3 Overcoming Androcentric Economic Ethics

The discussion about economic ethics restarted in the 1980s after a long period of silencing critical normative questions in positivist science and mathematical modeling. The different approaches to economic ethics can be classified as corrective, functionalist, and integrative (Ulrich 2008: 79–110; see also Nelson 2006).⁷ They will be scrutinized in the following from a critical feminist and reflexive perspective.

Corrective economic ethics is based on a two-world conception of supposed value-free economics on the one side and an external ethics on the other. The economy is understood as “an area still untouched by ethics into which morality must first be introduced . . . as a complement or corrective” (Ulrich 2008: 79). Ulrich’s main objection to this approach is directed against “the abandonment of reflection [*Reflexionsstopp*] in the face of the given ‘conditions of the market economy’ and the economic understanding of rationality as such” (ibid.: 87). Analogously, a corrective approach takes gender inequalities in the market economy as exceptions that could be avoided by the application of a gender-aware ethics. When injustice becomes too obvious, this ethics is applied to the (value-free) economy. The main limitation of this approach is the unawareness that the economic system itself is based on gender inequalities.

Functionalist economic ethics presumes that market competition leads to better results than the morality of human beings, so that “the normative content lies concealed solely in the logic of the ideal market economy system and the corresponding economic rationality” (ibid.: 97). This approach makes a *Reflexionsstopp* in front of the limitations of the underlying normative individualism and the social constructedness of individual preferences (ibid.: 99). The main limitations of functional economic ethics from a feminist perspective are the assumptions that economic rationality is taken as gender-neutral, that preferences are understood as given and not as cultural constructions, and that markets are trusted to solve gender inequalities (because otherwise there would be the threat of a loss of wealth). However, *homo economicus* as a gender-neutral utility-maximizer is not a suitable model to explain, for instance, why anyone would be interested in doing unpaid work.

Integrative economic ethics scrutinizes the normative foundations of economic rationality and reflects on them critically. Therefore, Ulrich distinguishes three general systematic tasks of integrative economic ethics (ibid.: 110, 269–71). The first is to make visible the normative implications of economic rationality. The second is “the development of guiding ideas of rational economic activity from the lifeworld perspective”

⁷ With the focus on feminism in social ethics, Sabine Plonz (2006: 25–6) distinguishes *supplementary ethics* (adding women to the discussed topics), *separation ethics* (women’s topics as separate topics), and the *feminist re-orientation of social ethics*.

(ibid.: 269). The central aim of economic activity is seen in “the service to life” (Rich 1990: 22) with two dimensions. One is the question of meaning or the ethics of the good life, asking which values should be created, and the other the question of legitimation or the ethics of justice asking for whom the values should be created (Ulrich 2008: 185–268). The third task is “the determination of the ‘sites’ of the systematic mediation between moral claims and the functional conditions of the economic system” (ibid.: 269). As possible sites of socio-economic responsibility, Ulrich (ibid.: part IV) works out regulatory ethics at the structural level, corporate ethics at the organizational level, and economic citizens’ ethics at the individual level. Integrative economic ethics has been concentrated on the sphere of paid work, on market processes, and enterprises, and the unpaid (household) economy is widely ignored. However, integrative economic ethics can be developed further from a feminist critical perspective to overcome the *Reflexionsstopp* concerning gender as analytical category in the economy, to include a critique of gender norms and gender relations underlying economic thinking and acting and making visible the institutionalized gender regimes (Knobloch 2015).

12.3 FEMINIST ECONOMIC ETHICS: METHOD, SUBJECT MATTER, AND MODEL OF ACTION

This section is based on several distinctions nested into each other. In developing a feminist critical reflexive *method* of economics, the three tasks of integrative economic ethics are scrutinized from a gender perspective with the third task distinguishing the three levels of economic ethics mentioned above (Section 12.3.1). With the help of the four-sector model—market, state, nonprofits, and households—the feminist critical reflection of the *subject matter* of economics leads to an economics of paid and unpaid work (Section 12.3.2). By searching for a universal and sustaining *model of action* in everyday life, the motives of caring and providing economic activities and their integration are examined on the three levels (Section 12.3.3).

12.3.1 Methodological Approach: Feminist Critical Reflection

Economic theories are neither value-free nor gender-neutral, but are based on often implicit gender norms at the individual level, gender relations at the organizational level, and gender regimes at the structural level. Since the early 1990s, I have been developing a feminist approach to economic ethics using a critical reflexive method to both make explicit and critically reflect the normative and gender-specific foundations of economic thinking and acting. In “A Different Economic Ethics?” (1993), I transferred the debate on ethics of care versus ethics of justice into economic ethics. This economic-ethical approach, which reflects the material foundations of economy and society, can be

explained as “the (re)introduction of the ethical perspective in looking at the foundations on which industrial economies and the theoretical thinking about them are based. Revealing the often taken for granted foundations is looking at economics from an ethical perspective” (Jochimsen and Knobloch 1997: 108; see also Plonz 2006, 2011; Nelson 1996: 59).

In what follows, I look more closely at the three systematic tasks of integrative economic ethics from a gender perspective.⁸ The first task can be developed further to make visible the traditional gender norms and, thereby, to overcome stereotypes, the underlying gender relations in market and household economics, as well as the gender regimes of economic systems. The aim is to create a feminist critical reflexive method without abandoning reflection concerning the category gender on the three distinguished levels.

From a gender perspective, the second task of integrative economic ethics includes the (good) lives of women in the question of meaning, and gender equality and women’s rights into the question of legitimation. The aim is to integrate the two dimensions of the just and the good in a non-androcentric way, for which the capabilities approach is a good example. With their capabilities approach, Amartya Sen (1999) and Martha Nussbaum (2000) refine the idea of social primary goods with its focus on income and wealth that John Rawls outlined in *A Theory of Justice* (1971). Fundamental to their capabilities approach is the distinction between what a person is actually able to do (*functionings*) and the real opportunities a person has (*capabilities*). Sen’s attempt to integrate the question of meaning and the question of legitimation is the idea of positive freedom as freedom “to do things that a person has reason to value” (1999: 56). Nussbaum works out a list of central capabilities of *all* human beings—with explicit attention to female capabilities—and understands them as “fundamental entitlements of citizens, all necessary for a decent and dignified human life” (2006: 166). Other capabilities approaches are either created within this framework, as in Robeyns (2003), or developed independently, as in Ulrich (2008: 247–52). Today, many feminist economists base their economic theories on the capabilities approach (such as Benería et al. 2016; Knobloch 2014).⁹

Looking at the third task of integrative economic ethics from a gender perspective leads us to develop a feminist economic ethics at three levels: the structural, the organizational, and the individual.

12.3.1.1 *Structural Level*

On the structural level, a regulatory ethics scrutinizes the normative principles and objectives behind the regulation of the economic framework. Two important concepts of regulatory policy are neoliberalism and ordoliberalism, both of which recognize the

⁸ In developing a feminist-critical social ethics, Sabine Plonz distinguishes two steps: “As a first step the implicit ethics of capitalist structural dynamics has to be uncovered... A second step would be that a feminist critique of political economy not only unmasks the dominant, practically and theoretically effective, wrong ethics gradually showing alternatives that are oriented towards the human scale. It would also have to make transparent its own theories of the good life and right action as well as its norms, social goals, anthropological theorems, and it needs to reflect how they become effective” (2011: 377–8, own translation).

⁹ See the chapter by Binder and Robeyns in this volume for more on the capabilities approach.

necessity of shaping the regulatory framework, but for different reasons. The aim of neoliberalism in shaping the regulatory framework is to facilitate competition; in contrast, ordoliberalism is focused primarily on shaping the regulatory framework to enable a good and just life of all people and only secondarily on enabling competition. Therefore, ordoliberalism is characterized by a *two-stage conception* of primary vital policy and subordinate competition policy (Ulrich 2008: 326–41).

Both stages need to be studied from a gender perspective. To make sure that vital policy is oriented towards the just and good life for *all* people, feminist scholars have pointed out that it has to address gender equity as well as justice concerning care and provisioning. Competition policy from a gender perspective has to be aware of the different work and life contexts of men and women, especially the greater responsibility of many women for unpaid work and the patriarchal dividend for many men. Regulatory ethics is also necessary in the global context. Naila Kabeer gives a good example to show the scope of the task of making visible the gender arrangements of welfare and care regimes on the one hand and to shape them on the other:

The struggle for labor standards needs to be broadened and made more inclusive by transforming itself into a struggle for a universal ‘social floor,’ so that all workers, men as well as women, urban as well as rural, formal as well as informal, in work and without it, are able to organize for their other rights without fear of jeopardizing their means of survival. (2004: 28)

12.3.1.2 *Organizational Level*

Because the economy is not only the market economy, ethics at the organizational level does not only concern corporate ethics including business or management ethics, but the ethics of all other economic organizations like households, state companies, and nonprofit organizations (NPOs) as well.

Business forms range from small and middle enterprises (SME) to global corporations, and from family businesses to stock companies. The aim of *corporate ethics*, which goes beyond corporate sustainable responsibility (CSR), is understood as “the unreserved critical examination of ethical aspects that might deserve priority over the pursuit of profit” (Ulrich 2008: 397). In increasing order of the degree “to which the inherent value of ethical points of view and the need to overcome the profit principle are recognized” (ibid.: 398), four types of corporate ethics are to be distinguished: *instrumentalist* (ethics as a means of profit-making), *charitable* (donations after profit-making), *corrective* (situative limitations in profit-making), and *integrative* (critical reflection of profit-making) (ibid.: 398–408). Integrative corporate ethics involves a two-stage conception of corporate responsibility: in the first stage, companies search for meaningful and legitimate value creation; and in the second stage, they take their share of responsibility for shaping the regulatory framework to facilitate meaningful and legitimate value creation (ibid.: 408–18).

Julie A. Nelson (2003, 2006) has revealed the androcentric bias of business ethics. Furthermore, the analysis of the distinctions of corporate ethics from a gender perspective

would allow for examination of how gender questions are handled in the different types of corporate ethics and how helpful the distinction of the two stages for a feminist corporate ethics would be. That would mean that in the first stage, companies have the responsibility for gender equality in their business, including equal opportunities, equal pay, and the reconciliation of paid and unpaid work. In the second stage, companies take their share of responsibility for scrutinizing the underlying gender norms, relations, and regimes, such as the glass ceiling or the glass escalator, the business relevant care gap, and “the assumption that workers have someone else at home taking care of the household” (Elson and Gideon 2005: 26).

Important economic organizations also include private households, ranging today from single households or father-mother-child families to community households or extended patchwork families. *Household ethics* is not limited to the ethics of household management, but looks also at the responsibilities that all household members take around organizing a household (see Arn 2000). Furthermore, the basics of economic rationality as well as the basics of a just and social life can be learned from family care and housework (see Okin 1989).

Analogous to corporate ethics, a two-stage conception of household ethics can be outlined. In the first stage, household ethics includes a critical reflection on needs and wants, the just distribution of unpaid work, organization of additional paid work inside the household, the creation of good working conditions for domestic workers, sustainable consumption, and more. In the second stage, households take their share of responsibility in shaping the economic framework to overcome the inequalities and privileges of the distribution of unpaid housework and care work, especially the patriarchal dividend and other unjustified privileges, and bring ideas of new and more equal distribution between the household members to the political debate (see Addabbo et al. 2010; Cooke 2010).

Feminist organizational ethics has to look at (paid and unpaid) work performed not only in companies and households, but in *and for* all other economic organizations as well.

12.3.1.3 *Individual Level*

Feminist economic ethics at the individual level understands citizens not as autonomous but as dependent persons and recognizes female economic citizens as full members and care citizenship as an important part, if not the starting point. Besides the share of civic responsibility as economic citizens, the critical loyalty as organizational citizens, and the responsibility as critical investors and consumers (Ulrich 2008: 288–314), it is fundamental for every society that citizens in addition assume responsibility for the individually and socially necessary share of unpaid work (not only) in the household focused on provisioning (Knobloch 2015), including consumption work understood as the work related to sustainable consumption, from information acquisition to waste disposal.

Another challenge to feminist critical reflexive economic ethics is the inclusion of the (new) ethical debates in queer and postcolonial studies (Section 12.1.3). On this basis, early thoughts about queer and postcolonial economic ethics can be found in the literature.

The task of *queer economics* is defined “as the examination of and response to the effects of heteronormativity both on economic outcomes and on economics as a discipline” (Jacobsen and Zeller 2008: 2; see also Gibson-Graham 1996, 2006). Heteronormativity can be persistent in feminist economics, too, but more radical approaches help to overcome this heteronormative thought (Badgett 1995; Danby 2007), such as examining the division of paid and unpaid labor in queer households (see Goldberg 2013) or in asking: “How do people develop desires about the manner in which their community, society and world is organized without being led only by the (hetero-) normative standards of given regimes of normalcy?” (Engel 2009: 159, own translation; for the global perspective see Dhawan et al. 2015). The ethical (re)turn in *feminist postcolonial economics* helps us to better understand “that some cultures have been ideologically devalued by the metaphysics of modernism in a parallel manner to the way in which women have been devalued in economics” (Zein-Elabdin 2003: 332). It is an ongoing task to develop further feminist critical thinking including queer and feminist postcolonial economic ethics.

12.3.2 Reflected Subject Matter: Paid and Unpaid Work

On the basis of the critical method outlined above, feminist economic ethics is able to reflect on the subject matter of economics. To avoid the abandonment of reflection concerning the material preconditions of market economy, feminist economic ethics includes all kinds of unpaid work usually grouped into housework, care work, and voluntary work. In the following, I look more closely at the definition of the subject matter of three important feminist economic approaches: critical household economics, the subsistence approach, and the economic theory of paid and unpaid work.

12.3.2.1 *Critical Household Economics*

In ancient Greece, economy was understood as the art of managing a household, but neither Xenophon or Aristotle, nor the paterfamilias literature later, critically reflects the “natural” distribution of labor determining the mode of living together in that household. With the industrial revolution and the developing market economy, the economic thought of the classical economists concentrated on the political economy neglecting the economy (of the household). Out of the variety of different tasks of a household, market economists were solely interested in consumption. *Home economics* of the nineteenth and early twentieth century has shown the importance of household production and its importance for the well-being of the household members (Beecher 1842; Reid 1934). *New Home Economics* (NHE) refers to these ideas, but uses only the very limited assumptions of neoclassical economics. Therefore, when it comes to the question of meaning it sticks to taken as given preferences of traditional households (see Section 12.2.2).

A *critical household economics* makes visible the underlying normativity of the assumed household model(s) applicable for the whole range of households. Today, the male breadwinner and female housewife model is replaced by several other models with different patterns of distribution of paid and unpaid work. But this does not mean that

they overcome the model's gendered distribution of labor and underlying gender norms. Based on Badgett (1995), Fraser (1997), Razavi (2007), and others, the list below gives an overview of possible models of the distribution of work within households.

1. *Male breadwinner and female housewife model*
Man as full-time earner, woman as full-time provider/carer (unpaid).
2. *Modified breadwinner and carer model*
Man as full-time earner, woman as short part-time earner and long part-time provider/carer with some support provided primarily by kin.
3. *Gendered earner model*
Man as full-time earner, woman as long part-time earner and short part-time provider/carer with a lot of support provided by kin and the market, public sector, and nonprofit sector.
4. *Universal earner model*
Adult members of a household as full-time earners; most of the previously unpaid provisioning and care work done in the family becomes paid work done in the market, public sector, and nonprofit sector.
5. *Universal earner and carer model...*
 - a) *of a two adult household*: Both adult members as part-time earners and part-time providers/carers (as early and as long in life as possible) with some support provided by the market, public sector, and nonprofit sector.
 - b) *of a single or lone parent household*: Adult member as part-time earner and part-time provider/carer (as early and as long in life as possible), lone parent with a lot of support of state benefits.
 - c) *of all other households*: All adult members as part-time earners and part-time providers/carers (as early and as long in life as possible) with some support provided by the market, public sector, and nonprofit sector.

Only in model 5 are paid and unpaid work equally distributed, but this model is not (yet) the norm of the provisioning and care regimes. Not reflected in this list are the necessary changes in the job markets, such as the commodification of care in model 3 needing people in the paid sectors doing the care work (Chorus 2013) and the questions around global households and housework abroad (Safri and Graham 2010).

12.3.2.2 *The Subsistence Approach*

Second-wave feminism has claimed the importance of housework for every society, even capitalist societies. Inspired by Rosa Luxemburg's *The Accumulation of Capital* (1913), Maria Mies, Veronika Bennholdt-Thomsen, and Claudia von Werlhof developed the so-called *subsistence approach* in the late 1970s, early 1980s (see, for instance, Bennholdt-Thomsen et al. 1992; Mies 1986). The subsistence approach has a global perspective with an emphasis on the global south. The work of all women, paid labor

included, is treated as if it were housework and, therefore, if not for free, then cheaply available—a phenomenon called *housewifization*.

Thereby, subsistence is understood as being able to provide for oneself and having enough to live on, and is connected with self-sufficiency and needs orientation. Until the 1950s, most people lived in a subsistence-oriented manner, cultivating and producing many necessities by themselves. Examples are vegetable gardens, small animal farming, preserving vegetables and fruits, knitting and crocheting, sewing and patching clothes, and crafting and repairing. Only since then has subsistence been increasingly displaced through mass production. But initiatives such as the do-it-yourself movement, urban gardening, repair cafés, and the like show that there is a rediscovery of subsistence in the twenty-first century (Baier 2010).

12.3.2.3 *The Economic Theory of Paid and Unpaid Work*

There is general consensus today that feminist economics has to go beyond a “Magic-Table Economy” (Knobloch 2002, 2016) that solely studies the market sector and ignores who prepares the meals, does the dishes, and so on. An *economic theory of paid and unpaid work*, in contrast, aims to include unpaid work, systematically looking not only into the market sector, but into the household sector, the public sector, and the nonprofit sector as well. Work, paid and unpaid, is done in companies, households, the state, and NPOs (Elson 2000; Razavi 2007; Knobloch 2016). By focusing on all four sectors of the economy and the shifting processes between them (shown in Figure 12.1), the twofold shifting process between paid and unpaid work comes to the fore. It consists on the one hand of shifts from the market, public sector, and nonprofit sector to the household sector, better known as *decommodification* or *demonetarization*. On the other hand, it consists of shifts from the unpaid work of the household sector to the other three sectors, better known as *commodification* or *monetarization*. Monetizing processes can also take place inside the households with (often poorly) paid persons doing provisioning and care work in other people’s homes.

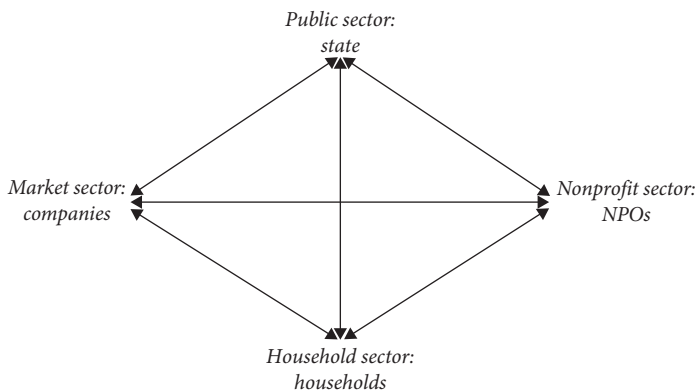


FIGURE 12.1 Shifting processes of work in the four sector-model.

An economic theory of paid and unpaid work is able to recognize that when unpaid work is turned into paid work, the gross domestic product (GDP) increases with no real increase in provisioning services—a phenomenon named *pseudo growth* (Knobloch 2015: 273). Furthermore, it is able to scrutinize “how economic policies which purport to improve the efficiency of resource use may instead shift costs from the paid economy to the unpaid economy” (Elson and Gideon 2005: 25–6).

12.3.3 Sustaining Economic Activities: Caring Provisioning

Feminist economic ethics scrutinizes human agency models looking for a model that is sustaining and gender-just. To this end, two important models of action in feminist economics, provisioning and caring, are outlined below, and then integrated.

In the early 1990s, Julie A. Nelson defined feminist economics as focussed “on the provisioning of human life, that is, on the commodities and processes necessary to human survival” (1993: 32, see also 1996: 36). Starting from this consideration, several authors have outlined an economics of provision or social provisioning, such as Marilyn Power (2004) and Marvin T. Brown (2010). The model of action contained therein is that of *provisioning work*. Its special characteristics are best described by Neysmith et al.:

Through provisioning we expand the meaning of work to include all activities that secure resources and provide the necessities of life to those for whom one has relationships of responsibility. This definition speaks to a range of activities that are never finished, must be performed regularly, and require energy and attention. These activities cannot be isolated or separated from the context of social relationships because provisioning consists of those daily activities performed to ensure the survival and well-being of oneself and others. Both the activities and the relationships may be voluntary or prescribed. The point is that the activities are necessary; without them, people would not survive. (2012: 4)

Provisioning work contributes significantly to individual and societal well-being. To learn more about this contribution, a feminist economics of provisioning has to distinguish between “to provide” and “to be provided” with the necessary goods and services; and to examine who does how much provisioning work for whom and who needs how much provisioning work from whom; at the three levels, four sectors, and two stages of the sectors’ organizations. Until today, provisioning work in public and private companies, NPOs, and households are institutionalized along the gender category. A critical reflexive ethics of sustainable provisioning makes visible the implicit heteronormativity of provisioning regimes, including underlying gender norms of individual provisioning as well as the gender-specific division of social provisioning, and gives orientation to what adequate individual and social provisioning and gender-just provisioning regimes could mean. Analogous to the concept of food sovereignty, the idea of provisioning sovereignty could be developed directed to rendering all people capable to provide for themselves, and to being provided for if needed, including the shifting of the perspective from private provisioning to public entitlement when necessary.

Referring to the debates on care ethics, considerations about *care economics* started in the early 1990s. Jochimsen and Knobloch (1997) based their idea of a caring economy on the importance of ethics and sustainable development, the integration of paid and unpaid work, and the principles of household economics. They conclude:

A whole economy organized according to the principles of caring activities [in an earlier version: according to the principles of housekeeping] would be a caring economy. In such a caring economy the satisfaction of the existing, material and non-material, basic needs would take priority over the production of new material goods. (1997: 111)

Against this background, caring economy is understood as a general economic theory learning from housework and care work.

Since the 2000s, care economics is one of the major topics in feminist economics. Its main assumptions are that every human being needs *care work* (paid and unpaid) to survive and that the effective provision of caring activities is a precondition of individual and societal well-being. On the individual level, care economics asks who does which forms of care work for whom and, thereby, makes visible and critically reflects the underlying gender norms. On the organizational level, it is shown in which sectors care work is offered under which conditions, thereby making visible and critically reflecting the underlying gender relations. On the structural level, it is worked out which kinds of caring activities social and economic systems presuppose and how the gender norms and gender relations around care work are shaped through the underlying gender regime (see Knobloch 2013a: 10–11). Furthermore, it is asked how the societal organization of care work, that is, the care regime, is and should be shaped, especially but not exclusively the care provision for children and the elderly (Daly 2001).

Starting her conceptualization from the provision of caring activities for dependents, such as young children and frail elderly persons, Maren A. Jochimsen (2003a, 2003b) has developed a theoretical concept called *Careful Economics* that is sensitive to asymmetries and dependencies as well as the concomitant underlying power structures. Reversing the fundamental assumptions of autonomy and independence in neoclassical economics by supposing limited autonomy and dependency as decisive assumptions, Jochimsen outlines an economic theory of careful provisioning that reaches beyond caring situations for dependents as “limits to autonomy in human life seem to be more the rule, and independent autonomy more the exception” (2003b: 243).

Nancy Folbre (2001) outlines an institutional approach to care economics, Mascha Madörin (2006) a feminist macroeconomics of paid and unpaid care work, and Shahra Razavi (2007) a political and social economy of care. With the help of these approaches focused on the structural level, it is possible to make explicit the often implicit normativity of care regimes and to analyze them from a gender perspective. This may be illustrated by the following examples and ideas: in the traditional breadwinner-housewife model, women (as “breadservers”) supply a significant majority of unpaid care work and are in return financially provided through the “breadwinners.” Thus, the care regime is mainly private. Besides the advantages of care migration for the migrants and their home

countries, the moral harm for the migrants and their families has to be recognized (Kittay 2014). In comparison to industrial goods, person-related services become more and more expensive due to a limited increase in labor productivity (Donath 2000; Madörin 2011).

Based on care economics, a feminist economic ethics can give orientation to good care at the individual level, to decent working conditions at the organizational level, and to care justice at the structural level, as well as to the sustainability of these principles at all three levels.

The concept of *caring provisioning* integrates the insights of provisioning and care economics and combines their models of action (Knobloch 2013b). The aim of a *feminist economics of caring provisioning* is to learn what is needed to perform caring provisioning activities to an adequate extent, of high quality, and under good working conditions, without reinforcing social inequalities, and to valorize unpaid provisioning and care work without further monetization and exploitation. A *feminist ethics of caring provisioning* could work with the distinguished levels and sectors to again critically reflect the normative foundations. At the individual level, the task is to examine the gender norms and the heteronormativity behind caring provisioning activities ranging from subsistence production to sustainable consumption. At the organizational level, the gendered distribution of caring provisioning activities comes to the fore including the question of who consumes and who produces most of the caring provisioning work in which organization. At the structural level, the materialization of social inequalities in caring provisioning regimes can be made visible. In strengthening the idea of care and provisioning rights and duties, a feminist ethics of caring provisioning can give orientation for vital and competition policy to overcome the inequalities of the current gender regimes and to examine the gender regimes around old and new institutional arrangements of caring provisioning, such as basic income and time banking. Linked to this is the awareness that caring provisioning is always a balancing act between under- and over-supply, fulfillment and sacrifice.

12.4 CONCLUSION

Feminist economic ethics uses a critical reflexive method to make visible the normative foundations of economic thought and action; defines a subject matter that includes all societal necessary activities in four sectors; and outlines an agency model that is directed at caring provisioning at three levels. Against the background of persistent socioeconomic gender inequalities and global care crises, a feminist economic ethics reveals privileges and power structures and gives orientation for striving to shape the necessary economic structures. To this purpose, feminist economic ethics requires a sound understanding of the underlying gender norms and gender relations in economic systems and provisioning policies, as well as of the embeddedness of market systems and neoliberal policies in gender and care regimes. In times of globalization and digitization, it can give orientation to transform society and the economy in the direction of provisioning sovereignty and

care justice for all people, while remaining aware that some of the necessary elements and perspectives are still missing.

The main topics of feminist economic ethics, accordingly, are not only women in the national and global paid labor markets and their working conditions, gender pay gaps, and equal opportunities for men and women in all economic sectors, but also the gendered distribution of paid and unpaid work, gender care gaps and care migration, living wages, and basic social security provision. A feminist economic ethics is also illuminating for the analysis of current topics like degrowth, basic income, commons, the digital revolution, and microfinance from a critical reflexive gender perspective, always including unpaid work and its distribution.

REFERENCES

- Addabbo, Tindara, Marie-Pierre Arrizabalaga, Cristina Borderías, and Alastair Owens (eds.). 2010. *Gender Inequalities, Households and the Production of Well-Being in Modern Europe*. Farnham: Ashgate.
- Anderson, Elizabeth. 1993. *Value in Ethics and Economics*. Cambridge, MA: Harvard University Press.
- Arn, Christof. 2000. *HausArbeitsEthik. Strukturelle Probleme und Handlungsmöglichkeiten rund um die Haus- und Familienarbeit in sozioethischer Perspektive*. Zurich: Rüegger.
- Badgett, M. V. Lee. 1995. "Gender, Sexuality, and Sexual Orientation: All in the Feminist Family?" *Feminist Economics* 1 (1): 121–39.
- Baier, Andrea. 2010. "Subsistenzansatz. Von der Hausarbeitsdebatte zur 'Bielefelder Subsistenzperspektive.'" In Ruth Becker and Beate Kortendiek (eds.), *Handbuch Frauen- und Geschlechterforschung. Theorie, Methoden, Empirie*. 3rd edn. (Wiesbaden: VS Verlag für Sozialwissenschaften), pp. 75–80.
- Bauhardt, Christine. 2017. "Living in a Material World: Entwurf einer queer-feministischen Ökonomie." *Gender. Zeitschrift für Geschlecht, Kultur und Gesellschaft/Journal for Gender, Culture and Society* 9: 99–114.
- Beauvoir, Simone de. 1949. *Le Deuxième Sexe*. 2 vols. Paris: Gallimard.
- Beecher, Catherine E. 1842. *A Treatise on Domestic Economy for the Use of Young Ladies at Home and at School*. 2nd edn. Boston, MA: T. H. Webb.
- Benería, Lourdes, Günseli Berik, and Maria S. Floro. 2016. *Gender, Development, and Globalization: Economics as if All People Mattered*. 2nd edn. London: Routledge.
- Benhabib, Seyla. 1992. *Situating the Self: Gender, Community and Postmodernism in Contemporary Ethics*. New York: Routledge.
- Bennholdt-Thomsen, Veronika, Maria Mies, and Claudia von Werlhof. 1992. *Frauen, die letzte Kolonie. Zur Hausfrauisierung der Arbeit*. 3rd edn. Zurich: Rotpunktverlag.
- Brown, Marvin T. 2010. *The New Economics of Provision*. Cambridge: Cambridge University Press.
- Butler, Judith. 2004. *Undoing Gender*. London: Routledge.
- Butler, Judith. 2005. *Giving an Account of Oneself*. New York: Fordham University Press.
- Charusheela, S. 2004. "Postcolonial Thought, Postmodernism, and Economics: Questions of Ontology and Ethics." In S. Charusheela and Eiman O. Zein-Elabdin (eds.), *Postcolonialism Meets Economics* (London: Routledge), pp. 40–58.

- Chorus, Silke. 2013. *Care-Ökonomie im Postfordismus. Perspektiven einer integralen Ökonomie-Theorie*. Münster: Westfälisches Dampfboot.
- Cooke, Lynn Prince. 2010. "The Politics of Housework." In Judith Treas and Sonja Drobnič (eds.), *Dividing the Domestic. Men, Women, and Household Work in Cross-National Perspective* (Stanford, CA: Stanford University Press), pp. 59–78.
- Daly, Mary. 2001. "Care Policies in Western Europe." In Mary Daly (ed), *Care Work. The Quest for Security* (Geneva: International Labor Office), pp. 33–55.
- Danby, Colin. 2007. "Political Economy and the Closet: Heteronormativity in Feminist Economics." *Feminist Economics* 13 (2): 29–53.
- Dhawan, Nikita, Antke Engel, Christoph F. E. Holzhey, and Volker Woltersdorff (eds.). 2015. *Global Justice and Desire. Queering Economy*. London: Routledge.
- Donath, Susan. 2000. "The Other Economy: A Suggestion for a Distinctively Feminist Economics." *Feminist Economics* 6 (1): 115–23.
- Elson, Diane (ed). 2000. *Progress of the World's Women 2000*. UNIFEM Biennial Report. New York: United Nations Development Fund for Women.
- Elson, Diane, and Jasmine Gideon. 2005. "Organising for Women's Economic and Social Rights: How Useful is the International Covenant on Economic, Social and Cultural Rights." *Bulletin Texte* 29/30: 14–30.
- Engel, Antke. 2009. *Bilder von Sexualität und Ökonomie. Queere kulturelle Politiken im Neoliberalismus*. Bielefeld: transcript.
- Ferber, Marianne A., and Julie A. Nelson (eds.). 1993. *Beyond Economic Man: Feminist Theory and Economics*. Chicago, IL: University of Chicago Press.
- Ferber, Marianne A., and Julie A. Nelson (eds.). 2003. *Feminist Economics Today: Beyond Economic Man*. Chicago, IL: University of Chicago Press.
- Folbre, Nancy. 2001. *The Invisible Heart: Economics and Family Values*. New York: New Press.
- Fraser, Nancy. 1997. *Justice Interruptus: Critical Reflections on the "Postsocialist" Condition*. New York: Routledge.
- Garber, Marjorie, Beatrice Hanssen, and Rebecca L. Walkowitz (eds.). 2000. *The Turn to Ethics*. New York: Routledge.
- Gibson-Graham, J. K. 1996. *The End of Capitalism (As We Knew It)*. Oxford: Blackwell.
- Gibson-Graham, J. K. 2006. *A Postcapitalist Politics*. Minneapolis, MN: University of Minnesota Press.
- Gilligan, Carol. 1982. *In a Different Voice: Psychological Theory and Women's Development*. Cambridge, MA: Harvard University Press.
- Gilman, Charlotte Perkins. 1911. *The Man-Made World or, Our Andro-centric Culture*. New York: Charlton.
- Goldberg, Abbie E. 2013. "'Doing' and 'Undoing' Gender: The Meaning and Division of Housework in Same-Sex Couples." *Journal of Family Theory and Review* 5: 85–104.
- Hanlon, Niall. 2012. *Masculinities, Care and Equality: Identity and Nurture in Men's Lives*. Basingstoke, UK: Palgrave Macmillan.
- Hartmann, Heidi I. 1979. "The Unhappy Marriage of Marxism and Feminism: Towards a More Progressive Union." *Capital and Class* 3 (2): 1–33.
- Heimbach-Steins, Marianne. 2009. "Human Rights – Whose Benefit? Critical Reflections on the Androcentric Structure of Human Rights and its Consequences on the Social Participation of Women." In Beat Sitter-Liver (ed), *Universality: From Theory to Practice* (Fribourg: Academic Press), pp. 229–57.
- Hoppe, Hella. 2002. *Feministische Ökonomik. Gender in Wirtschaftstheorien und ihren Methoden*. Berlin: Edition Sigma.

- Irigaray, Luce. 1984. *Éthique de la Différence Sexuelle*. Paris: Éditions de Minuit.
- Jacobsen, Joyce, and Adam Zeller (eds.). 2008. *Queer Economics: A Reader*. London: Routledge.
- Jennings, Ann L. 1993. "Public or Private? Institutional Economics and Feminism." In Marianne A. Ferber and Julie A. Nelson (eds.), *Beyond Economic Man: Feminist Theory and Economics* (Chicago, IL: University of Chicago Press), pp. 111–29.
- Jochimsen, Maren A. 2003a. *Careful Economics: Integrating Caring Activities and Economic Science*. Dordrecht et al: Kluwer.
- Jochimsen, Maren A. 2003b. "Integrating Vulnerability: On the Impact of Caring on Economic Theorizing." In Drucilla K. Barker and Edith Kuiper (eds.), *Toward a Feminist Philosophy of Economics* (London: Routledge), pp. 231–46.
- Jochimsen, Maren A., and Ulrike Knobloch. 1997. "Making the Hidden Visible: The Importance of Caring Activities and their Principles for any Economy." *Ecological Economics* 20: 107–12.
- Kabeer, Naila. 2004. "Globalization, Labor Standards, and Women's Rights: Dilemma of Collective (In)Action in an Interdependent World." *Feminist Economics* 10 (1): 3–35.
- Kant, Immanuel. 1785. *Groundwork of the Metaphysics of Morals*. In *Practical Philosophy*, Mary J. Gregor (trans. and ed.) (Cambridge: Cambridge University Press, 1996), pp. 37–108.
- Kittay, Eva Feder. 1999. *Love's Labor: Essays on Women, Equality, and Dependency*. New York: Routledge.
- Kittay, Eva Feder. 2014. "The Moral Harm of Migrant Carework: Realizing a Global Right to Care." In Alison M. Jaggar (ed), *Gender and Global Justice* (Cambridge: Polity Press), pp. 62–84.
- Knijn, Trudie, and Monique Kremer. 1997. "Gender and the Caring Dimension of Welfare States: Toward Inclusive Citizenship." *Social Politics* 4: 328–61.
- Knobloch, Ulrike. 1993. "Eine andere Wirtschaftsethik? Die Bedeutung der Frauenfrage für die Begründung einer grundlagenkritischen Wirtschaftsethik." *Beiträge und Berichte* 59. St. Gallen: Institut für Wirtschaftsethik.
- Knobloch, Ulrike. 2002. "Kooperation in der feministischen Wirtschaftsethik." In Adelheid Biesecker, Wolfram Elsner, and Klaus Grenzdörffer (eds.), *Kooperation und interaktives Lernen in der Ökonomie* (Frankfurt am Main et al: Peter Lang), pp. 151–71.
- Knobloch, Ulrike. 2013a. "Sorgeökonomie als kritische Wirtschaftstheorie des Sorgens." *Denknetz Jahrbuch* 2013: 9–23.
- Knobloch, Ulrike. 2013b. "Versorgen – Fürsorgen – Vorsorgen. Normative Grundlagen einer Sorgeökonomie als allgemeine Wirtschaftstheorie und die Ethik des Vorsorgenden Wirtschaftens." In Netzwerk Vorsorgendes Wirtschaften (ed.), *Wege Vorsorgenden Wirtschaftens* (Marburg: Metropolis), pp. 17–42.
- Knobloch, Ulrike. 2014. "Questioning the Gender-Based Division of Labour. The Contribution of the Capabilities Approach to Feminist Economics." In Flavio Comim and Martha C. Nussbaum (eds.), *Capabilities, Gender, Equality: Towards Fundamental Entitlements* (Cambridge: Cambridge University Press), pp. 195–214.
- Knobloch, Ulrike. 2015. "Kritische Wirtschaftsethik aus Geschlechterperspektive." In Thomas Beschorner, Peter Ulrich, and Florian Wettstein (eds.), *St. Galler Wirtschaftsethik. Programmatik, Positionen, Perspektiven* (Marburg: Metropolis), pp. 263–82.
- Knobloch, Ulrike. 2016. "Jonglieren mit Zeiten: Wirtschaftstheorie der bezahlten und unbezahlten Arbeit." In Monica Budowski, Ulrike Knobloch, and Michael Nollert (eds.), *Unbezahlt und dennoch Arbeit* (Zurich: Seismo Verlag), pp. 25–54.
- Komlosy, Andrea. 2014. *Arbeit. Eine globalhistorische Perspektive. 13. bis 21. Jahrhundert*. Vienna: Promedia.
- Kuiper, Edith, and Jolande Sap (with Susan Feiner, Notburga Ott, and Zafiriz Tzannatos) (eds.). 1995. *Out of the Margin: Feminist Perspectives on Economics*. London: Routledge.

- Luxemburg, Rosa. 1913. *Die Akkumulation des Kapitals. Ein Beitrag zur ökonomischen Erklärung des Imperialismus*. Berlin: Vorwärts.
- Madörin, Mascha. 2006. "Plädoyer für eine eigenständige Theorie der Care-Ökonomie." In Torsten Niechoj and Marco Tullney (eds.), *Geschlechterverhältnisse in der Ökonomie* (Marburg: Metropolis), pp. 277–97.
- Madörin, Mascha. 2011. "Das Auseinanderdriften der Arbeitsproduktivitäten. Eine feministische Sicht." *Denknetz Jahrbuch* 2011: 56–70.
- Marx, Karl. 1867. *Capital: A Critique of Political Economy*. Moscow: Progress Publishers (1887 edition).
- Mayhew, Anne. 1999. "Value." In Janice Peterson and Margeret Lewis (eds.), *The Elgar Companion to Feminist Economics* (Cheltenham, UK: Edward Elgar), pp. 732–7.
- Mies, Maria. 1986. *Patriarchy and Accumulation on a World Scale: Women in the International Division of Labor*. London: Zed Books.
- Mohanty, Chandra Talpade. 2003. *Feminism Without Borders: Decolonizing Theory, Practicing Solidarity*. Durham, NC: Duke University Press.
- Nelson, Julie A. 1993. "The Study of Choice or the Study of Provisioning? Gender and the Definition of Economics." In Marianne A. Ferber and Julie A. Nelson (eds.), *Beyond Economic Man: Feminist Theory and Economics* (Chicago, IL: University of Chicago Press), pp. 23–36.
- Nelson, Julie A. 1996. *Feminism, Objectivity and Economics*. London: Routledge.
- Nelson, Julie A. 2003. "Separative and Soluble Ethics: Androcentric Bias and Business Ethics." In Marianne A. Ferber and Julie A. Nelson (eds.), *Feminist Economics Today: Beyond Economic Man* (Chicago: University of Chicago Press), pp. 81–99.
- Nelson, Julie A. 2006. *Economics for Humans*. Chicago, IL: University of Chicago Press.
- Neysmith, Sheila M., Marge Reitsma-Street, Stephanie Baker Collins, and Elaine Porter. 2012. *Beyond Caring Labour to Provisioning Work*. Toronto: University of Toronto Press.
- Noddings, Nel. 1986. *Caring: A Feminine Approach to Ethics and Moral Education*. Berkeley, CA: University of California Press.
- Nussbaum, Martha C. 2000. *Women and Human Development: The Capabilities Approach*. Cambridge: Cambridge University Press.
- Nussbaum, Martha C. 2006. *Frontiers of Justice: Disability, Nationality, Species Membership*. Cambridge, MA: Belknap Press of Harvard University Press.
- Okin, Susan Moller. 1989. *Justice, Gender, and the Family*. New York: Basic Books.
- Pearse, Rebecca, and Raewyn Connell. 2016. "Gender Norms and the Economy: Insights from Social Research." *Feminist Economics* 22 (1): 30–53.
- Plonz, Sabine. 2006. *Arbeit, Soziale Marktwirtschaft und Gesellschaft. Studienbuch Feministische Sozialethik*. Koblenz: Neukirchener Verlag.
- Plonz, Sabine. 2011. "Mehrwert und menschliches Maß. Zur ethischen Bedeutung der feministisch-ökonomischen Care-Debatte." *Das Argument* 292: 365–80.
- Polanyi, Karl. 1944. *The Great Transformation: The Political and Economic Origins of Our Time*. Boston, MA: Beacon Press (2001 edition).
- Power, Marilyn. 2004. "Social Provisioning as a Starting Point for Feminist Economics." *Feminist Economics* 10 (3): 3–19.
- Rawls, John. 1971. *A Theory of Justice*. Cambridge, MA: Harvard University Press.
- Razavi, Shahra. 2007. "The Political and Social Economy of Care in a Development Context: Conceptual Issues." Research Questions and Policy Options, *Gender and Development Programme Paper*, Number 3, Geneva: UNRISD.
- Reid, Margaret G. 1934. *Economics of Household Production*. New York: Wiley.

- Rich, Arthur. 1990. *Wirtschaftsethik II: Marktwirtschaft, Planwirtschaft, Weltwirtschaft aus sozialetischer Sicht*. Gütersloh: Gütersloher Verlagshaus Gerd Mohn.
- Rich, Arthur. 1991. *Wirtschaftsethik I: Grundlagen in theologischer Perspektive*. 4th edn. Gütersloh: Gütersloher Verlagshaus Gerd Mohn.
- Robeyns, Ingrid. 2003. "Sen's Capability Approach and Gender Inequality: Selecting Relevant Capabilities." *Feminist Economics* 9 (2–3): 61–92.
- Robinson, Fiona. 1999. *Globalizing Care: Ethics, Feminist Theory, and International Relations*. Boulder, CO: Westview Press.
- Robinson, Fiona. 2013. "Global Care Ethics: Beyond Distribution, Beyond Justice." *Journal of Global Ethics* 9: 131–43.
- Safri, Maliha, and Julie Graham. 2010. "The Global Household: Toward a Feminist Postcapitalist International Political Economy." *Signs* 36: 99–125.
- Schnabl, Christa. 2005. *Gerecht sorgen. Grundlagen einer sozialetischen Theorie der Fürsorge*. Fribourg: Academic Press.
- Sen, Amartya. 1999. *Development as Freedom*. Oxford: Oxford University Press.
- Sevenhuijsen, Selma. 1998. *Citizenship and the Ethics of Care: Feminist Considerations on Justice, Morality and Politics*. London: Routledge.
- Smith, Adam. 1759. *The Theory of Moral Sentiments*. D. D. Raphael and A. L. Macfie (eds.). Indianapolis, IN: Liberty Classics (1982 edition).
- Smith, Adam. 1776. *An Inquiry into the Nature and Causes of the Wealth of Nations*. 2 vols. R. H. Campbell, A. S. Skinner, and W. B. Todd (eds.). Indianapolis, IN: Liberty Classics (1981 edition).
- Steuart, James. 1767. *An Inquiry into the Principles of Political Economy: Being an Essay on the Science of Domestic Policy in Free Nations*. 2 vols. London: A. Millar and T. Cadell.
- Thiem, Annika. 2008. *Unbecoming Subjects: Judith Butler, Moral Philosophy, and Critical Responsibility*. New York: Fordham University Press.
- Todorova, Zdravka. 2014. "Consumption as a Social Process." *Journal of Economic Issues* 48: 663–78.
- Tronto, Joan C. 1993. *Moral Boundaries: A Political Argument for an Ethic of Care*. New York: Routledge.
- Tronto, Joan C. 2013. *Caring Democracy: Markets, Equality and Justice*. New York: New York University Press.
- Ulrich, Peter. 2008. *Integrative Economic Ethics: Foundations of a Civilized Market Economy*. Cambridge: Cambridge University Press.
- Veblen, Thorstein B. 1899. *The Theory of the Leisure Class: An Economic Study of Institutions*. New York: Macmillan.
- Veblen, Thorstein B. 1904. *The Theory of Business Enterprise*. New York: Charles Scribner's Sons.
- Veblen, Thorstein B. 1914. *The Instinct of Workmanship and the State of the Industrial Arts*. New York: Macmillan.
- Warner, Michael. 1991. "Introduction: Fear of a Queer Planet." *Social Text* 29: 3–17.
- Wærness, Kari. 1984. "The Rationality of Caring." *Economic and Industrial Democracy* 5: 185–212.
- Zein-Elabdin, Eiman O. 2003. "The Difficulty of a Feminist Economics." In Drucilla K. Barker and Edith Kuiper (eds.), *Toward a Feminist Philosophy of Economics* (London: Routledge), pp. 321–38.

CHAPTER 13

ECONOMY AND CULTURE

The Importance of Sense-Making

ARJO KLAMER

“CULTURE matters in the economy.” Such a statement is to be expected in a conversation among anthropologists and sociologists, but it is becoming increasingly current in the company of economists. Deirdre McCloskey, the economic historian, argues in her trilogy on bourgeois virtues (2007, 2011, 2016) that cultural factors, not economic ones, account for the surge of the Dutch economy in the early seventeenth century. For the Nobel Prize-winning economist Amartya Sen (2004), it is not even a question whether culture matters in economic processes, so he focuses on *how* culture matters, which is also the position the economists Beugelsdijk and Maseland (2011) take. The Austrian economist Virgil Storr wants to show in his work that “culture shapes economic behavior and outcomes” (2003: 1). Economists like Alberto Alesina and Paola Giuliano (2015) have surveyed the research on the impact of culture on institutions, including economic ones. And urged by the sociologist Richard Florida (2002), economists such as Andy Pratt (2009) are increasingly interested in exploring the role of creatives in urban regeneration.

In the meantime, cultural economics has become a burgeoning field of its own. In one approach, it encompasses the application of economic methods to the study of the arts and creative industries (Frey 2003; Throsby 2001; Towse 2013). In another, it refers to the investigation of the economic impact of cultural factors in subjects such as development (Rao and Walton 2004), finance (Karolyi 2016), institutions (Alesina and Giuliano 2015), and organizations (Cameron and Quinn 1999; Hofstede 2003; Magala 2009). When we add all the humanists who approach economic subjects with cultural concepts, and the anthropologists and sociologists who do the same, the academic literature on the interaction between the economy and culture is getting quite crowded (Leemans and Johannes 2013; Woodmansee and Osteen 1999; Walzer 1983; Sandel 2012).

This chapter aims to provide an understanding of the way social scientists, and economists in particular, are making sense of the interaction between “culture” and “economy.” The ambition is not to give a complete survey; there are simply too many

contributions to cover them all. The intended contribution is to introduce an ordering of all the arguments by way of a model of five spheres that emanates from the so-called *value based approach to the economy* (Klamer 2017).

By way of this model, we will see why the mainstream of economists remains oblivious to the work on the relationship between the economy and culture: the methodological framing of economics stands in the way. If economists were to overcome their mainly methodological objections, they may recognize the potential of adding the cultural dimension. Doing so, however, may affect the practice of economists in fundamental ways. It may, for example, change the dominant methodology of “model building,” motivate the relaxation or even abandonment of “rational choice” as the main heuristic, change the definition of the subject of economics, lead to a reappraisal of the notion of “values,” and inspire greater recognition of the importance of cultural policies. More importantly, bringing in the cultural dimension highlights *sense-making activity* as a crucial and substantive part of making a life and a society.

13.1 A MODEL OF FIVE SPHERES, EACH WITH A DIFFERENT LOGIC

Consider that the subject of economics is how people make a life, not only materially, but also immaterially. Let us assume that people seek a good life and want to be living in a good society. To those ends they have a family, get jobs, start a company, go out shopping, buy things like a mango or a car, have a house (and a home), make friends, join clubs, become politically active, go on trips, play music, paint, have hobbies, and so on. By bringing in the notion of culture, we add the consideration of what makes a meaningful life. Being able to make sense of all they do to make a living, people may be able to realize a meaningful life. That is why they engage in all kinds of sense-making practices, like religion, art, science, like the reading and writing of books, surfing the web, the watching and making of movies, and so on. Thus, we perceive them as trying to realize what is important to them, that is, to valorize their values (Klamer 2017: ch. 4).

The model that I present here suggests that there are basically five spheres in which individuals and groups of people operate to seek a good life and valorize their values. Each sphere has its own logic, that is, a particular way of doing things, its own institutions, characteristic values and rhetoric—a pertinent manner of speaking (Klamer 2017).

The idea that people operate in different spheres is quite common, although not so among economists. Karl Polanyi (1944) distinguishes four spheres: the sphere of exchange, redistribution, reciprocity, and of house holding. The philosopher Jürgen Habermas (1987) developed the distinction between the life world—where we interact socially and informally—and the world of systems—where we act formally in the context of organizations. The sociologist Viviane Zelizer (2005) has analyzed the differences between the payments that people make in the sphere of exchange, in a governmental

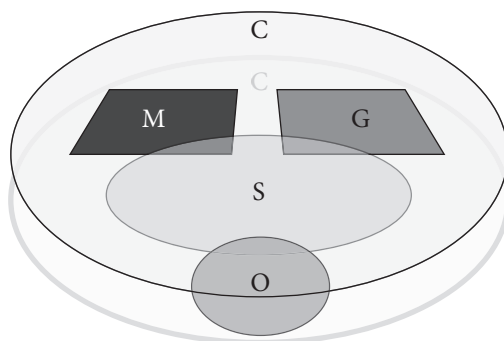


FIGURE 13.1 The model of five spheres.

sphere and in a social or intimate sphere. The economist Albert Hirschman (2007) made the (among economists) famous distinction between the “exit,” “voice,” and “loyalty” options to get things done. The sociologists Boltanski et al. (2006) distinguish six worlds of value regimes (on grounds that are not all that clear).

All these distinctions inform the model of the five spheres, with the added dimension of the cultural sphere. (See Figure 13.1.) People start the making of a life at home (O stands for *oikos*, Greek for home). It is where they grow up, take their first steps, have parents or other elders who feed and shelter them, learn to play, and socialize. Sooner or later they most likely will leave their first home in order to establish another. Usually with a partner they will run a household, sometimes with kids of their own. At home, they generate essentials like companionship (unless they choose to remain single), meals, shelter, relaxation, entertainment, health, education (of themselves and the children), intimacy, perhaps love, and care. At home, people exercise Hirschman’s option of loyalty, and home is part of the lifeworld of Habermas.

When, in his *Politics*, Aristotle began thinking about something like an economy he focused on the household. Economy to him was about *oikos*—home or village—and *nomein*—to distribute or manage, or the running of the household (broadly defined). That starting point still makes sense, at least if we think from the perspective of people seeking a good life. Accordingly, the O is at the bottom of the diagram in Figure 13.1; it is the base as the anthropologist Gudeman (2008) calls it.

When people get outside their *oikos*, they step into the public space. They attend schools for their education, socialize with people in the neighborhood, form friendships, join clubs, become politically active, or other things you do in the social sphere, S, which stands for all social activities that people undertake trying to make a life. In S, people form relationships of all kinds; reputations are made and broken; people have a status, or a social role, such as police officer, teacher, social activist, politician, or parent; form communities; develop collegiality and trust with others (or fight and try to become powerful). Hirschman would say that in the social sphere people exercise the option of voice, although the option of loyalty still applies as well. The social sphere is also part of the life world.

And yes, to make a life, people have to deal with and operate in markets. In the market sphere, M, they exchange with other people and organizations. They go to shops to buy things that they need in their *oikos* or for their social activities. In the so-called labor market they offer their time, knowledge, and talents in exchange for amounts of money that they need, in turn, to purchase things. Here the option of exit applies: when you do not like the offering of one, you go to another. Economists like Milton Friedman, Friedrich von Hayek, and Deirdre McCloskey are eager to herald the amazing results of the market logic, and rightly so. However, it is only one of the five spheres that we need to valorize our values.

Maybe even more important is the sphere of governance (G). Organizations play a large role in the making of a living. When people seek work in exchange for a salary, they need to deal with organizations, with firms in particular. They seek employment in an organization or set up an organization themselves. Within organizations they produce and create a great variety of goods and services for other organizations and for people in need of those for their livelihood. The exchange with people takes place through organizations like shops (both physically and online). They also need to deal with an organization called the government, by respecting its laws and regulations, paying taxes and receiving benefits. Civil servants do their work in the government.

When they work within organizations or deal with organizations, people operate in the sphere of governance in the diagram. In this sphere, they are subject to the rules and laws of organizations and have to deal with bureaucracy with its forms, procedures, criteria, accounting, planning, and budgeting. G stands for the world of systems as Habermas states it. (The question is whether the market sphere also belongs to this world; I would think so as market logic, like that of governmental, is anonymous and impersonal.)

In the cultural sphere C, people make sense of their world and life and weave their “web of significance.” In this sphere, they operate on a meta level: they reflect on what they do in all the other spheres, develop frames for thinking about them (by way of the science of economics, for example), figure out what they have in common with other people, and try to add meaning to their life by way of religion, art, literature, movies, television, social media, music, design, stories, rituals, poetry, philosophy, and science. This is also the sphere in which they reflect on what are the right things to do, or what is appropriate and morally just behavior.

One important outcome of this model is the insight that people engage in distinctly different practices when they make a life. Practices are ongoing activities. People do not go to the grocery store once, but do so periodically. Work is practice and so is playing a sport. Some of us practice science, or music, or do things regularly with friends. The notion of practice encourages us to think beyond a mere transaction in the market sphere, or something static like a society. We are made to think of friendship, entrepreneurship, family, bureaucracy, and trade as practices, as ongoing activities in which we participate and to which we contribute. We are led to ask what these practices are good for. When we watch people practice the buying of groceries, we can guess that they are intended to facilitate other practices, such as family or friendship. Why then do people engage in art, or economics?

Note that what is logical in one sphere is not necessarily so in another sphere. You do not ask your partner for money for a service delivered and you do not expect a grocery owner to give you a bottle of wine without paying for it, even if you are friendly with him and you asked him graciously. The logic of M just does not work in the O and S, and the logic of S does not work in M. Likewise, you will not appeal to the love and care of the civil servant when you try to get your unemployment benefit. You could do so, but civil servants would risk losing their job for giving benefits on such grounds. And it would be strange to draw a little economic model to explain what is happening when someone is trying to make a deal with you. Such an action would make sense in a classroom but not on a trading floor. Of course, there are overlaps. In reality, spheres can get mixed up; we operate in several at the same time, like when we socialize with the shopkeeper or gallery owner. Even so, it is important to start the analysis with the distinctiveness of each of these five spheres impressed on our mind.¹

13.2 THE NOTION OF CULTURE

The inclusion of the cultural sphere in the model is critical for the examination of the interaction between culture and economy. The convention in economics is to ignore the cultural dimension altogether, the arguments of prominent economists such as Sen, Alesina, and McCloskey notwithstanding. For that matter, the spheres of the *oikos* and the social are not in the conventional economic picture either, as that picture focuses on what happens in the spheres of M and G, and on M in particular.

Because this article focuses on the interaction between culture and economy, or, better, the way economists and other social scientists make sense of this interaction, we need to make sense of what we mean when we speak about “culture” and “economy.” Both concepts are problematic, so it turns out.²

“Culture” does not appear in standard economic textbooks whereas the concept overwhelms texts in anthropology. The concept is contested and subject to a range of interpretations. Clifford Geertz, a well-known anthropologist, is often cited with his definition:

The concept of culture I espouse . . . is essentially a semiotic one. Believing, with Max Weber, that man is an animal suspended in webs of significance he himself has spun, I take culture to be those webs, and the analysis of it to be therefore not an experimental analysis in search of laws but an interpretative one in search of meaning. It is explication I am after. (1974: 4–5)

If we follow Geertz in his focus on culture as webs of significance, we can distinguish three senses of culture (Klamer 2017: 8–9).

¹ See Klamer (2017: chs 9–10) for a more extensive analysis.

² For an exhaustive discussion of the definitional issues see Beugelsdijk and Maseland (2011). However, they do not recognize artistic, religious, and scientific practices as a separate category as I do here (as C₃).

C1 is culture in the anthropological sense: it stands for all the values, history, and identities that a group of people share, and with which they distinguish themselves from other people. In this sense, an economics department can have a distinctive culture, just as Sicily has, or Belgium, or people who are into square dancing, or people who make up an organization.

C2 is culture in the sense of civilization. This is the culture that comprises (cultural) heritage of a (usually extensive) group of people, its accumulated works of science, art, literature, architecture, design, and fashion, but also the manners and institutions that such a group heeds (Arnold 1915; Elias 2000; Scruton 2010; Huntington 1996; Hall 1998). Accordingly, we speak of a Western and an Eastern civilization, or the ancient Roman, Greek and Egyptian civilizations. (Civilizations can perish, as these last three showed.)

C3 is culture in the sense of artistic, religious, scientific, or other symbolic practices, that is, cognitive or expressive practices by which people try to make sense of their lives and their world. All these practices have in common that they generate symbolic expressions, including concepts, images, rituals, and icons. We may also consider all sense-making practices that gravitate around products of cultural industries such as music, musicals, movies, television programs, design, and (creative) crafts.

The cultures define more or less the identity that people seek and claim. In C3 people derive an identity from being engaged in certain practices. They are “artists” when practicing art and “scientists” when practicing science. In C2 they can identify themselves with their civilization, claiming to be “European,” “modernist,” or “a tribal person” (when tribal life is seen as a civilization). And the umbrella of C1 gives people national, regional, city, organizational, or communitarian identities.

By differentiating culture in these three meanings, we are able to distinguish different interactions between culture and economy. Is culture in the anthropological sense (C1) the issue or an artistic practice (C3)? If we are looking at the culture of organizations, then we consider the sphere of governance, G. And it becomes conceivable that identities, certainly of the C1 types, may affect decisions, as they appeared to do when the British voted for Brexit in 2014. Such far-reaching decisions usually expose sharply contrasting ways of sense-making. A large group of British voters wanted to associate themselves with “Europa,” and the values of innovation and international collaboration, whereas the majority put their identification with Britain first. Finally, when the economy is the subject of concern, we most likely look for the interaction between C1 and M, that is, between culture in the anthropological sense and whatever takes place in the sphere of the market.

13.3 ECONOMICS AS A CULTURAL PRACTICE: DOES IT IMPACT THE ECONOMY?

The exploration starts with a consideration of the practice of “us,” the author and the implied readers of this chapter. How would that practice figure in the constellation of

the five spheres? And how to make sense of its interactions with and influence on the other spheres? Does economics matter?

In line with the definition of C₃, the practice of economics, economic sociology, economic anthropology, or cultural economics as a science, are all a way of making sense of the phenomenon that we call the economy, and therefore is a cultural activity. The accumulated knowledge of economics is part of C₂, civilization. (I leave out the consideration that the community of academic economists constitute a culture in the sense of C₁, even though that is an interesting topic).

To start with the obvious, the characterization of a set of human activities as constituting “the economy” is a sense-making activity, too. Economists may have a clear sense of what constitutes an economy, but most other people do not. The average Jane and Joe will think of “money” when asked about the economy. Few will think of “the production, distribution, and consumption of commodities,” as economists may articulate their subject; fewer even will think of scarcity and rational choice, as the core of economists will formulate their approach to the economy. Then again, not many economists may realize that their characterizations of economics and the economy are drawn from an essay of Lionel Robbins published in 1932, and a textbook by Paul Samuelson published first in 1947. The point is that the economy is not what it is, but what economists and whoever else make it to be on the metalevel of C₃.³

Accordingly, with the interaction between culture and economy at issue, we are led to wonder how economics as a sense-making activity (and part of C₃) interacts with the activities in M, G, O, C, and S (in Figure 13.1). In what way does the theorizing of economists, all the work they have done during the last century and a half, affect what people do, or how they make sense of what they do? Or is the influence the other way? May the actual economy, such as the material conditions of production, determine what economists think (as is the argument in Marxist sense-making)?

Such questions are barely asked, at least not in the literature. Economists do not seem to worry much about the impact of their work. They hardly ever consider the relevance of their knowledge for the agents whose behavior they study. For economists who believe in the efficiency of free markets, that makes sense. After all, what can scientific knowledge contribute to the functioning in efficient markets? Yet, economists who believe in the effectiveness of economic policies should be concerned. Callon (1998) introduced the notion of *performative economics* to investigate the way economics performs outside of academia (see also MacKenzy et al. 2007). It is a good set-up for a range of case studies of, for example, the application of game theory in the design of auctions, or the impact of finance economics on instruments in the financial market (Faulhaber and Baumol 1988). The relevance of economics for everyday life, or, for that matter, for the design of policy, still needs to be determined, though.

I should also mention the attempt of so-called cultural economists to make sense of the world of the arts (C₃) in economic terms (Throsby 2001). They analyze the arts in

³ For a good discussion of the development of economic thinking in relationship to the notion of culture, see Beugelsdijk and Maseland (2011: chs 2–4).

terms of demand and supply, productivity (with one result being the Baumol disease, the observation that productivity gains in the arts lack those in other sectors), labor markets, and public goods. Doubts are increasing whether such a sense-making makes sufficient sense of the art world (Frey 2003; Klammer 1996; Throsby 2001). All this work continues to support the instrumentalist approach of mainstream economics (Taylor 1991; Tinbergen 1975). In such an approach, economic research is supposed to deliver results that can be used as input for rational decisions on actual economic policies. (Should governments subsidize the arts, or not?) Preferences and values are considered to be beyond the reach of economic analysis and are presumed to be given. What is left is the attention to the instruments of policy, such as governmental expenditures, taxes, interest rates, exchange rates, regulations, and laws.

The problem, however, is that the impact of this instrumental approach is far from obvious and direct. In various investigations, the relationship between outcomes of economic research and actual policy was found to be complex and often absent (Coats 2001; Cordes et al. 1993; Klammer and Meehan 1999). Considering that economic research often generates various and at times conflicting outcomes on the same subject, we cannot be too surprised. The question then is what is the impact of economics? Does economics matter? Are things changing because of what economists think?

The lack of research on these questions is remarkable. A good reason is that determining the effect of ideas on action is hard to do. Ideas have the habit to travel in funny ways, sometimes taking years to affect what people do, getting altered in the process. When people cite Marx, Keynes, or Hayek to justify their actions, their students may doubt the accuracy and appropriateness of those citations. Where non-economists invoke economics when they speak of markets and money, economists may be shocked at the lack of understanding and the imprecision of their use of the terms. Moreover, the interaction between ideas and actions is hard to trace. People may believe one thing and do another. Even so, if the point of economics is to influence actions, like policies, the lack of research into its effectiveness remains remarkable, to put it mildly.

One point to make is that economics is rhetorical as McCloskey (1998), Klammer (2007), and Mirowski (2013) have pointed out. Persuasion is not a mechanical process where inputs of information transform in a predictable and traceable manner into knowledge that, in turn, produces specific actions or decisions. The process is messy and complex, involving time lags, an abundance of signals, miscommunication, and erratic moments of insights. As most readers of this chapter are educators, too, they may be aware of the complexity of learning, of the difficulties of conveying to freshmen the economic way of thinking, including the meanings of concepts (like elasticity, opportunity costs, depreciation, aggregate demand and so on), not to mention the challenge of getting them to make sense of real-life situations with these concepts. Learning is a messy and complex process.

Even so, if economics does not affect policy decisions in a predictable, mechanical way, it may indirectly affect decision-making processes by influencing the sense-making of the economy by policy makers. As Mirowski (2013), Foucault (2008), Zuidhof (2012), and Klammer (2007) have shown, policymakers and especially their advisors may adopt

concepts and some of the reasoning—that is, the rhetoric—of economists when they discuss the economy. Neoliberalism is defined by Foucault and Zuidhof as the application of market (M) logic in the governmental (G) sphere. It happens when civil servants stress the importance of competition, efficiency, creative incentives by varying prices and remunerations, demand-oriented programs, and people seeking welfare as customers. Then again, the influence of economic logic and rhetoric showed up already in the 1930s when civil servants began to copy their accounting, differentiating consumption from investment spending and stocks from flows. The impact of national income accounts on policy discourse is probably the best example of the persuasiveness of economists. How that accounting logic came to be so successful up until now is a story that still needs to be told. In another setting, we are led to wonder how and in what ways the writings of Karl Marx and Mao could be so persuasive that they seemed to inform the sense-making of nearly half the world population for almost a century.

13.4 THE METHODOLOGICAL ISSUE

The case of scientific practices illustrates the methodological problems that a researcher interested in the interactions between cultural and economic practices faces. How to determine causal effects as an instrumental approach calls for?⁴

We will see that one way of dealing with the challenge is to insist on the transformation of vague phenomena such as culture into variables that are measurable somehow and so can be inputs in a model. Such an analytical approach fits the methodology of model-building and conducting regression analyses that characterizes mainstream economics. Cultural factors are viewed as constraints for rational agents or as factors that influence their preferences, or they affect institutions such as norms or laws, which in turn constrain rational decisions (Alesina and Giuliano 2015).

Another approach follows interpretive or hermeneutic methods. Acknowledging that sense-making is part of making a life, researchers want to explore the meanings that certain cultural practices (such as economics) generate and whether such meanings can be detected in other practices. Such work is more about characterizing practices, about searching for patterns, and tracing influences and interactions. It makes use of surveys, ethnographic methods, historical studies, case studies, and thick descriptions, as Geertz (2005) characterized his anthropological method. Such methods prevail in anthropology, history, sociology to some extent, the humanities, and some economics. As we shall see below, Deirdre McCloskey and Virgil Storr are two economists who use historical and interpretive methods.

When the two approaches meet, conflict is imminent. Scientists are most passionate, in my experience, not in their politics but in their methodological commitment (Klamer

⁴ For further elaboration of the methodological issues see Beugelsdijk and Maseland (2011: ch. 5). Intriguing is the absence of a discussion of causality, which may be related to their ignoring the instrumental issue, that is, how consideration of cultural factors can affect (political) decisions.

1984, 2007). Harsh judgments of “the others” are common. The way to deal with the threat of conflict is to ignore the others. It shows in the remarkable fact that the protagonists of the one approach hardly cite the literature that represents the other approach. Their practices are strictly divided. (DiMaggio 1994: 29).

Another question concerns the objectives of the explorations we are going to discuss. What purposes do they serve? Is it to provide advice to policymakers, as is the (often implicit) objective of standard economics? Is it to inform a (different) world view? Does it want to influence the sense-making by the general public? Or are they merely scientific or intellectual pursuits, aimed at furthering the scientific or intellectual conversation?

13.5 CULTURAL PRACTICES INFLUENCING OR CAUSING ECONOMIC PRACTICES

The reference with which each discussion on the cultural impact on the economy starts is Max Weber. In his *Protestant Ethic and the Spirit of Capitalism* (1905), Weber famously argued that the emergence of Protestantism at the expense of the Roman Catholic church in Northern European countries in the sixteenth and seventeenth century brought about a revolution in values that proved to be conducive for capitalistic practices. Weber is explicit about not claiming to have a cause for the emergence of capitalism; Protestant ethics is a contributing factor. Referring especially to the teachings of Calvin, Weber notes that his values such as parsimony, planning, using your talents, and seeing a “calling” in work began to inform economic life in Northern European countries at that time.

The idea that religious teachings had an effect on daily life, including the practices directed at making a life, makes sense. It is especially plausible for the times that Weber explores, with the majority of people receiving weekly the religious teachings in church. When the message of the clerisy stresses non-worldly pursuits, such as asceticism, penance, and good deeds, and strenuously warns against profit-seeking and charging interest, it should not come as a surprise that people act accordingly, condemning bankers and commercial types while admiring monks and people working in poor houses. People who engage in Buddhist practices with the mantra of letting go of the ego, being respectful of all life, and practicing compassion, may find it difficult to boast of a pursuit of richness and fame (Brown 2017; Payutto 1998). When the Islamic clergy insists that, according to sharia, charging interest is wrong, it should not come as a surprise that Islamic banking practice differs from other banking practices (Visser 2013). When the Pope issues an encyclical, he hopes that it will affect the practices of his followers. The question is whether and to what extent these religious practices influence or contribute to the cultural environment in which people operate and what the impact is on their actions and practices.

Plausible as the Weber thesis may be, however, it proves to be subject to various criticisms. For example, it has been argued that the spirit of capitalism was first detectable in

Italian cities in a Roman Catholic environment a few centuries earlier. And how did the influence actually work? Protestantism was very much in force in seventeenth-century Scotland but did not generate capitalism as it did in the Dutch Republic (Herman 2001), not after Scotland came under English rule.

Just as problematic is the argument that Confucian values account for the emergence of the Japanese economy in the 1970s (Sen 2004). They may have played a role in fostering employee loyalty and the focus on long-term planning in Japanese companies, but why did they not in the preceding century when Confucian values were in force as well? It is the suggestion of a causal connection between religious or philosophical practices and behavior in the market and organizational spheres (M and G) that is problematic.

In her trilogy on bourgeois virtues, Deirdre McCloskey (2007, 2011, 2016) traces the emergence of egalitarian and liberal values in the Dutch Republic and other Northern European countries in the sixteenth and seventeenth centuries, and observes how they brought about a culture of honor for tradespeople and an appreciation of entrepreneurship, risk-taking, and innovation. It was in such a culture, she argues, that new financial and organizational forms became possible that enabled the Dutch to equip ships to sail to territories far away and return with lucrative goods. Based on her exhaustive historical research, McCloskey concludes that ideas, that is, sense-making activities (C₃), in a particular cultural setting (C₁), account for an unprecedented period of economic growth. She resists using the term of capitalism and prefers to speak of “trade-tested progress” (2016: 3). It is with a culture of betterment that the Northern European economies began to thrive, so she argues. A culture of entrepreneurship, trade, and betterment overtook the culture of aristocracy and clerisy, which promulgated the values of distinction, class, and honor on one side, and those of sacrifice, humility, and charity on the other. The latter culture was anti-commercial whereas the new culture made commerce look good.

McCloskey engages in a sense-making activity, directing her exploration to other scholars to impress on them the importance of virtue ethics. She joins a scholarly movement that is bringing about a revival of the teleological ethics that Aristotle so distinctly articulated in his *Nicomachean Ethics*. She moreover tries to revive interest in and respect for Adam Smith, the moral philosopher and author of *The Theory of Moral Sentiments* (1759). In this respect she is contributing to the scientific conversation (C₃), with a faint chance to affect activities in M, G, S, and O. Yet, as she makes painstakingly clear, her historical and philosophical (interpretive) study is also an advocacy for liberalism with the main message that the unleashing of market forces that the moral turn-around made possible is responsible for staggering improvements in economic welfare the world around. To convey that message, she actively seeks interactions in the S by presenting her arguments in public fora and appearing in various media. In this regard she reminds us of the role that Friedrich Hayek played. He, too, recognized the importance of moral beliefs and civilization (Dekker 2015), and he was a main advocate of the value of freedom and markets as its instrument.

Advocacy is also the aim of the contribution of Lawrence Harrison and Samuel Huntington in their edited book *Culture Matters* (2000), but in a different direction. They are especially keen on stressing that if certain cultures facilitate economic progress, other

cultures stand in its way. When coupled with the conviction that not much can be done about changing cultures, their position is a pessimistic stance towards programs to promote progress in the context of collectivist and traditional cultures. This is also the outcome of an often-cited economic analysis of the collectivistic cultures on economic growth by Avner Greif (1994), who, with a standard economic model and a regression analysis, concluded that individualist cultures do better economically than collectivist cultures. For his data he drew from the numerous cited statistical work of Geert Hofstede (2003), based on surveys among IBM employees across the world in the early 1970s.

In contrast, the work of Virgil Storr and Emily Chamlee-Wright is more interpretive and less political. Their main concern is the lack of recognition of the cultural factor in economics, and they operate mainly in C3. Like McCloskey they stress the positive role of culture, such as when it is conducive for entrepreneurship or plays a constructive role in the reconstruction of New Orleans after Hurricane Katrina, pointing out the role that churches and their pastors had in maintaining social cohesion and communal assistance (Chamlee-Wright 2015). Storr and Chamlee-Wright operate in the Austrian tradition and build on the work of Don Lavoie, who advocated the use of interpretive methods (Lavoie and Chamlee-Wright 2015). In that way they make sense of a complex economy in which people act with limited knowledge and a need to make sense of their actions and the world in which they act. According to them understanding of the cultural context is indispensable for such a sense-making.

Introducing culture in the economic conversation invites attention to values and qualities. As Rao and Walton purport, a “focus on culture is necessary to confront the difficult questions of *what* is valued in terms of well-being, who does the valuing, and why economic and social factors interact with culture to unequally allocate access to a good life” (2004: 4) Whereas standard economics accentuates the values in exchange, equating value with price (as in exchange value or monetary value), the values that a culture fosters are rather answers to the question “what is important?” McCloskey (2007) focuses on what people consider important in behavior, or their virtues. Anthropological and sociological studies as well as the value-based approach to economics (Klamer 2017) explore the values or qualities that people attribute to things, states of affairs, and practices. The argument here is that certain values, such as those expressed in a cultural identity, affect economic behavior.

In an extensive survey of the economic and sociological work done on the interaction between culture and economy, the sociologist Paul DiMaggio (1994) uses the organization of themes according to the standard economic model. In the sphere of production, he discusses the work on corporate culture, where the issue is the influence of C1 on G, the working of organizations. A great deal more work has been done since on the subject. The influence on organizational practices shows in the recognition of the role of values (Cameron and Quinn 1999) and the realization that cultural factors account for the failure of mergers and acquisitions (Weber 1996). In business economics the cultural argument has become generally accepted (Magala 2009); in standard economic studies of industrial organization or public administration it is still virtually non-existent (for reasons noted above). The same outcome applies to cultural studies of consumption.

Whereas the importance of the cultural factor in consumptive behavior is obvious to anthropologists and sociologists, economists remain oblivious, wedded as they seem to the assumption of exogenous preferences.

The anthropologist Mary Douglas observed the cultural factor in risk-taking attitudes, a conclusion that we also can draw from the statistical work of Hofstede (2003). MacKenzy (2008) observes the influence of culture in financial markets, and Rao and Walton (2004) collect studies that demonstrate the role of culture in economic development. All these studies fall into the genre “culture matters in economic practices.”

A different approach takes Alesina in his own work and that with Giuliano in a survey of research into the impact of culture (that he defines as a set of informal institutions) on formal institutions (to be identified in the M and G spheres in Figure 13.1). Alesina and Giuliano (2015) focus in their survey on research that uses a standard economic approach, rather than an interpretive one, in which culture is to be defined as a variable somehow, to be expressed in numbers in order to determine a quantifiable (if not causal) relation with other variables. Although DeMaggio also considers some such studies, there is almost no overlap in their surveys.

Alesina and Giuliano find a wide range of relationships in the literature. Shared values affect the attitude of groups of people towards redistribution of income, poverty, work, and female participation in the workforce. When trust is viewed as a cultural factor, research shows its impact on a variety of economic processes, like the effectiveness of regulation, altruistic behavior, and economic development. Trust proves also to be related to the strength of family ties—another cultural factor—with the effect that strong family ties account for low trust in civil society. Family values appear to account for the performance of family business and resistance to external authorities. And cultural factors tend to be responsible for the strength of democratic institutions, as well as adherence to property laws and other regulations.

Alesina and Giuliano add important caveats. Apart from the difficulties of defining and capturing a complex phenomenon like “culture” in a quantifiable variable—a necessary condition for their kind of research—the causality in the relations between cultural and economic factors proves to be highly dubious. The chain of reactions is hard to trace, and reciprocal influence is often likely. Causality, however, is crucial for application of the outcomes in decisions (especially political ones).

The instrumentalist approach poses one even more serious problem. Even if cultural factors are instrumental for, say, economic development, female participation in the labor force, or effectiveness of economic regulation, they are of no use because they cannot be changed, managed, or otherwise governed. Politicians and managers do not have knobs to change cultural factors to their liking as they may do with financial variables or formal institutions. Cultural factors appear to be like geographical or weather conditions: they do affect economic processes yet that is all we can say about them. There is not much to do about them (unless we can move our operation or our life to another place with another, more congenial culture).

What remains is one strand of research that argues quite effectively for an instrumental use of culture in economic policy. The literature on creative cities purports that artistic

practices (C₃) can regenerate urban areas, stimulate creative industries and boost (local) economies. The sociologist Richard Florida (2002) made a name for himself in local policy circles across the globe with his prediction that an influx of creative workers has a positive impact on local economies. Cultural economists and social geographers followed suit, with arguments on behalf of the formation of creative clusters, investments in cultural institutions (with the museum of Bilbao as the prime example), and the attention of an increasing share of creative industries in local economies (Pratt 2009; Scott 2004). The formula seemed to be too good to be true, though, and so it was. As Florida (2017) himself has come to acknowledge, the formula is far from universally valid, pointing at the relevance of other factors such as culture in the sense of C₁, and, more importantly, it engenders increasing inequality and neglected neighborhoods with all kinds of undesirable (social) consequences.

13.6 HOW ECONOMIC PRACTICES MAY AFFECT CULTURAL PRACTICES

When we consider the impact of culture on economic practices, we focus on the effects of shared thoughts, beliefs, and values on the realization of a good life. But the effect may also be in the reverse order. Concerted economic actions (and transactions), mainly in the spheres of M and G, may affect what we share with others in terms of what we think, believe, and value (in the sphere of C). When people join a company, they may assimilate its culture. When financial markets soar, people may become risk-takers. Chinese culture may adjust to the recent take-off of the Chinese economy. The emergence of competitive forces and a stimulating environment can make people appreciate entrepreneurship more.

Here, too, the establishment of causality is treacherous. Did the rise of bourgeois virtues account for the increase in trading by the Dutch in the sixteenth century or was it the other way around? McCloskey (2016) reports the rise of an appreciation for bourgeois virtues but does so with recognition of the complexity of factors at work at that time, and so refrains from making causal claims. Cultures are complex and so are economies. We can distinguish patterns and correlations, and maybe determine influences, but a first mover remains elusive.

Even so, Marxist analyses tend to stress the causal effect of economic practices. Marx already claimed that the mode of production, the base, causes consciousness and with that the superstructure of what we now would call culture: “It is not the consciousness of men that determines their existence, but their social existence that determines their consciousness” (1859: 2). Adorno and Horkheimer, representing the so-called Frankfurt School, argued that culture is profoundly affected and informed by the machinations of cultural industries. In pursuit of profits, those operating in the entertainment industry seek to form ever larger companies to gain ever more control over the content that

gets produced (Horkheimer and Adorno 1944). Those companies determine more or less what music people listen to, what media they consult, what movies they watch, and what they read. The sociologist Pierre Bourdieu (1977) argues that financial wealth determines the level and kind of education that people receive, and with that their “cultural capital,” that is, their ability to realize cultural values appropriate for the circles in which they move. One’s financial position supposedly informs one’s social and cultural position. “Consuming” art is, therefore, a way for the highly educated class to distinguish themselves from other people (recalling Veblen’s 1899 analysis of the leisure class.)

Marxist analysis is less in vogue nowadays. Most studies that suggest an economic impact on culture refrain from making causal claims. Robert Lane (1991), for example, analyses a market culture that comes about when markets make inroads in all spheres of life. Richard Titmus (1970) is famous for his study showing that when the market logic gets introduced in the giving of blood, people are less willing to donate. Jacques Olivier et al. (2014) have shown that when trade barriers go down between regions or countries, their cultures converge. Earlier, Emile Durkheim (1893) studied the impact of the increasing division of labor on society; Friedrich Engels (1884) argued that the capitalistic mode of production had a profound influence on family life with the institution of the nuclear family; George Simmel (1900) traced the social and cultural impact of the use of money; and Karl Polanyi (1944) is often cited for his claim that the emergence of markets upset the traditions and institutions of feudal society making for an entirely different society.

In another vein, the economist Bruno Frey (1997) argues that the societal encroachment of the logic of M and G crowds out social arrangements, and with that certain social values; his argument is similar to that of sociologist Viviane Zelizer (2005, 2010). Accordingly, when civil servants get used to buying services in the market (M logic), they may overlook initiatives of citizens (S logic), and when children get used to the care that governments provide for their ageing parents (G logic), they may depreciate the importance of their own responsibility (O logic). Stronger appreciation of M or G logic, therefore, may bring about a change in social values and with that a change in culture (C₁) and possibly sense-making activities (C₃). Then again, the question remains what brought about the stronger appreciation of markets or governmental control? Was it something that has to do with culture?

13.7 CONCLUSION: IT IS ABOUT SENSE-MAKING

There is a great deal to be learned from all this work on the interaction between culture and economy. It generates insights on all kinds of economic processes, makes us aware of the complexity of what we call “the economy,” “culture,” and their interaction. It is tempting, indeed, to admonish economists who ignore and overlook the cultural

dimension. Culture does matter. Cultural differences do account for different economic outcomes. Creative practices do contribute to urban regeneration. Introducing market or governmental logic in segments of society may affect the social infrastructure and does change cultural practices.

Yet, there is a nagging doubt that lingers when reviewing all the work on the interaction between culture and economy. Instrumentalists may wonder why bother about culture if the variable does not transform into an instrument with which politicians or organizations can engineer better outcomes. When scholars such as McCloskey conclude that a liberal culture was conducive for economic development, or when others argue that collective cultures are hindering such, politicians may wonder what follows. How should they consider the cultural consequences of their policy of liberalization and privatization? What follows when governments take over the control of elderly care? Such observations do not come with guides on how to promote a liberal culture, to transform a collectivist culture into an individualist one, or how to deal with the loss of social values due to new policies. There are no knobs to turn. It makes for an instrumentalist paralysis.

If instrumentalism, with its focus on quantification, is quite senseless when it comes to cultural processes, there are other means to bring about changes. These tend to be qualitative in kind. Education is one of such means, as academics know all too well. Education affects the knowledge, attitudes, and the values with which graduates enter the “real world.” Social movements can bring about value changes as well, and by spilling over in political movements, they bring about changes in policy. It is conceivable that increasing awareness of cultural factors may lead to a re-evaluation of cultural policy as one of the most important policies for all organizations. After all, that policy is about purpose, or what is most important to the people who have a stake in the organization (such as a social enterprise, a cultural organization, a local or national government).

The impact of sense-making activity on everyday practices does not seem to be on the mind of most researchers. Their work is about truth-finding and is directed at fellow scholars. Typical scholars want to account for processes as they actually happen. To accomplish this, they distance themselves from their subjects and heed a distant and neutral perspective. At least that is the intention. Whether it is the work that follows the standard economic approach, such as Alesina and Giuliano (2015), or the interpretive work of Storr, McCloskey, and Beugelsdijk, it all shows that culture matters in economic processes. It is a matter of truth. Cultural differences are important to reckon with when accounting for economic development, female participation in the labor force, risk-taking, entrepreneurship, savings, remuneration and so many other economic phenomena. Economic outcomes are clearly related to cultural processes (C₁ in particular, but also artistic and religious practices, that is, C₃). The possible impact of their own sense-making is not part of the discussion, though.

Even so, scholarly work is a cultural activity itself: it is a sense-making practice (C₃). The work that we covered turns out to be critical of the kind of sense-making that characterizes standard economics with its focus on market (M) logic and, to a lesser extent, governmental (G) logic. It tries to make people think of the cultural context in which transactions take place, of the values that a certain culture privileges, and how those values

affect (economic) behavior. The working with values appears to be especially relevant in social, *oikos*, and cultural practices. Because values and qualities are not quantifiable, and acquire interpretive work when people act upon them, the metaphor of a calculating individual that drives standard economics does not suffice. If so, the modeling strategy becomes dubious as well. Alesina and Giuliano (2015), both standard economists, acknowledge the limitations of such a strategy. The consequence might be that economists have to learn how to investigate qualities of all kinds, take into account values, and develop methods to monitor qualitative outcomes.

With the maturation of cultural sense-making in economics, the possibility of impact on the sense-making by governors of organizations, policymakers, and people at work and home arises. It may affect their web of significance, as Clifford Geertz would put it. Where standard economic sense-making suggests that values are generated by way of market exchange (M logic) and the instrumental logic of G, the sense-making that highlights the cultural dimension turns the attention to the spheres of the *oikos*, the social world, and cultural practices. It makes people aware of the importance of a purpose of one's actions and of being engaged in good practices, and to contribute to them. Such sense-making makes people consider their values other than monetary quantities, and sensitize them to the conflicting logics of the various spheres in which they need to operate to realize their values. Cultural sense-making gives another meaning to everyday actions by placing the substance of making a good life above the instrumental elements.

Accordingly, where standard economic reasoning is satisfied with registering the amount of money that people spend in stores as consumption, the value-based approach explores what a purchase is good for or what qualities or values it is supposed to help realize. When people attend a theatre to improve their acting skills, they do not "consume" a performance but contribute to their craftsmanship. People who shop may actually buy the means to contribute to family life. Furthermore, where standard economics appears to promote the image of people as egoistic and greedy beings, a value-based framework rather casts them as social and cultural beings, eager to realize social goods and cultural goods that add meaning to their life.

A value-based sense-making may already be reflected in the sense-making by people in organizations. Although neoliberal sense-making, with its emphasis on values of the market logic such as the discipline of the market, probably still prevails, interest in the role of values and, with that, culture appears to be increasing. It shows in the interest for social enterprises, the awareness of corporate social responsibility, the continuing emphasis on sustainable development, and in leadership instead of management (as leadership is about getting one's values clear and management is about the instrumental aspects of organizations).

The consequence of these movements is that the economist John Maynard Keynes may be right after all with his prophesy that he wrote in 1931:

I see us free, therefore, to return to some of the most sure and certain principles of religion and traditional virtue—that avarice is a vice, that the exaction of usury is a

misdemeanor, and the love of money is detestable, that those walk most truly in the paths of virtue and sane wisdom who take least thought for the morrow. We shall once more value ends above means and prefer the good to the useful. (371–2)

Keynes foresaw that people would be preoccupied with what is “foul” rather than “fair” for another 100 years. Were “culture” to become the mainstay in the sense-making of economic practices, and were people to acknowledge the substantive relevance of their social and cultural preoccupations, that time might approach sooner than Keynes could have imagined.

REFERENCES

- Alesina, Alberto, and Paola Giuliano. 2015. “Culture and Institutions.” *Journal of Economic Literature* 53: 898–914.
- Arnold, Matthew. 1915. *Essays in Criticism*. New York: A. L. Burt Company.
- Beugelsdijk, Sjoerd, and Robbert Maseland. 2011. *Culture in Economics*. Cambridge: Cambridge University Press.
- Boltanski, Luc, Laurent Thévenot, and Catherine Porter. 2006. *On Justification: Economies of Worth*. Princeton, NJ: Princeton University Press.
- Bourdieu, Pierre. 1977. *Outline of a Theory of Practice*. Cambridge: Cambridge University Press.
- Brown, Claire. 2017. *Buddhist Economics: An Enlightened Approach to the Dismal Science*. New York: Bloomsbury Press.
- Callon, Michael. 1998. *The Law of Markets*. London: Blackwell.
- Cameron, Kim, and Robert Quinn. 1999. *Diagnosing and Changing Organizational Culture: Based on the Competing Values Framework*. New York: Addison-Wesley.
- Chamlee-Wright, Emily. 2015. “Pastor Response in Post-Katrina New Orleans: Navigating the Cultural Economic Landscape.” In Laura E. Grube, and Virgil H. Storr (eds.), *Culture and Economic Action* (Cheltenham, UK: Edward Elgar), pp. 269–94.
- Coats, A.W. 2001. “The Role of Economists in Government and International Agencies: A Fresh Look at the Field.” *History of Economics Review* 34: 19–32.
- Cordes, Joseph, Arjo Klamer, and Tim C. Leonard. 1993. “Academic Rhetoric in the Policy Arena: The case of capital gains taxation.” *Eastern Economic Journal* 19: 459–79.
- Dekker, Erwin. 2015. *Hayek and the Role of Civilization*. Cambridge: Cambridge University Press.
- DiMaggio, Paul. 1994. “Culture and Economy.” In Neil Smelser and Richard Swedberg (eds.), *The Handbook of Economic Sociology* (Princeton, NJ: Princeton University Press), pp. 27–57.
- Durkheim, Emile. 1893. *Division of Labour in Society*. London: Free Press (1997 edition).
- Elias, Norbert. 2000. *The Civilizing Process*. Oxford: Wiley-Blackwell.
- Engels, Friedrich. 1884. “The Origin of the Family, Private Property, and the State.” In Karl Marx and Friedrich Engels, *Marx/Engels Selected Works* (Moscow: Progress Publishers, 1969), pp. 191–334.
- Faulhaber, Gerard. R., and William J. Baumol. 1988. “Economists as Innovators: Practical Products of Theoretical Research.” *Journal of Economic Literature* 26: 577–600.
- Florida, Richard. 2002. *The Rise of the Creative Class*. New York: Basic Books.
- Florida, Richard. 2017. *The New Urban Crisis*. London: OneWorld Publishers.
- Foucault, Michel. 2008. *The Birth of Biopolitics: Lectures at the Collège de France, 1978–1979*. Michael Senellart (trans.), Graham Burchell (ed.). New York: Palgrave MacMillan.

- Frey, Bruno. 1997. *Not Just For The Money*. Cheltenham, UK: Edward Elgar.
- Frey, Bruno. 2003. *Arts and Economics*. Berlin: Springer.
- Geertz, Clifford. 1974. *Myth, Symbol, and Culture*. New York: Norton.
- Geertz, Clifford. 2005. "Deep Play: Notes on the Balinese Cockfight." *Daedalus* 134 (4): 56–86.
- Greif, Avner. 1994. "Cultural Beliefs and the Organization of Society: A Historical and Theoretical Reflection on Collectivist and Individualist Societies." *Journal of Political Economy* 102: 912–950.
- Gudeman, Stephen. 2008. *Economy's Tension: The Dialectics of Community and Market*. Oxford: Berghahn Books.
- Habermas, Jürgen. 1987. *The Philosophical Discourse of Modernity: Twelve Lectures*. Frederick Lawrence (trans.). Cambridge: Polity Press.
- Hall, Peter. 1998. *Cities in Civilization*. London: Weidenfeld & Nicolson.
- Harrison, Lawrence E., and Samuel P. Huntington (eds.). 2000. *Culture Matters: How Values Shape Human Progress*. New York: Basic Books.
- Herman, Arthur. 2001. *How the Scots Invented the Modern World*. New York: Crown Publishers.
- Hirschman, Albert O. 2007. *Exit, Voice, and Loyalty: Responses To Decline in Firms, Organizations, and States*. Cambridge, MA: Harvard University Press.
- Hofstede, Geert. 2003. *Culture's Consequences: Comparing Values, Behaviors, Institutions and Organizations Across Nations*. 2nd ed. Thousand Oaks, CA: SAGE.
- Horkheimer, Max, and Theodor W. Adorno. 1944. *Dialectiek van de Verlichting Filosofische Fragmenten*. Michel van Nieuwstadt (trans.). Nijmegen: Uitgeverij Sun (1987 edition).
- Huntington, Samuel P. 1996. *Clash of Civilization and the Remaking of World Order*. New York: Simon and Schuster.
- Karolyi, George A. 2016. "The Gravity of Culture for Finance." *Journal of Corporate Finance* 41: 610–25.
- Keynes, John Maynard. 1931. *Essays in Persuasion*. New York: Norton & Company (1963 edition).
- Klamer, Arjo. 1984. *Conversations with Economists: New Classical Economists and Their Opponents*. Totowa, NY: Rowman and Littlefield.
- Klamer, Arjo. 1996. *The Value of Culture*. Amsterdam: Amsterdam University Press.
- Klamer, Arjo. 2007. *Speaking of Economics: How to Get in the Conversation*. New York: Routledge.
- Klamer, Arjo. 2017. *Doing the Right Thing: A Value Based Economy*. London: Ubiquity Press.
- Klamer, Arjo, and Jennifer Meehan. 1999. "The Crowding Out of Academic Economics: The Case of NAFTA." In Robert F. Garnett, Jr. (ed.), *What Do Economists Know? New Economics of Knowledge* (London: Routledge), pp. 65–85.
- Lane, Robert E. 1991. *The Market Experience*. Cambridge: Cambridge University Press.
- Lavoie, Don, and Emily Chamlee-Wright. 2015. "How Does Culture Influence Economic Development?" In Laura E. Grube, and Virgil H. Storr (eds.), *Culture and Economic Action* (Cheltenham, UK: Edward Elgar), pp. 147–79.
- Leemans, Inger, and Gert-Jan Johannes. 2013. *Worm en Donder*. Utrecht: Prometheus.
- MacKenzy, Donald. 2008. *Material Markets*. Oxford: Oxford University Press.
- MacKenzy, Donald, Fabian Muniesa, and Lucia Siu. 2007. *Do Economists Make Markets? On the Performativity of Economics*. Princeton, NJ: Princeton University Press.
- Magala, Slawek. 2009. *The Management of Meaning in Organizations*. Basingstoke: Palgrave Macmillan.
- Marx, Karl. 1859. *A Contribution to a Critique of Political Economy*. Moscow: Progress Publishers, 1977 edition. Available at <https://www.marxists.org/archive/marx/works/1859/critique-pol-economy/preface.htm>

- McCloskey, Deirdre N. 1998. *The Rhetoric of Economics*. 2nd edn. Madison, WI: University of Wisconsin Press.
- McCloskey, Deirdre N. 2007. *The Bourgeois Virtues: Ethics for an Age of Commerce*. Chicago, IL: University of Chicago Press.
- McCloskey, Deirdre N. 2011. *Bourgeois Dignity: Why Economics Can't Explain the Modern World*. Chicago, IL: University of Chicago Press.
- McCloskey, Deirdre N. 2016. *Bourgeois Equality: How Ideas, Not Capital or Institutions, Enriched the World*. Chicago, IL: University of Chicago Press.
- Mirowski, Philip. 2013. *Never Let a Serious Crisis Go to Waste: How Neoliberalism Survived the Financial Meltdown*. London: Verso.
- Olivier, Jacques, Nicolas Maestre, Mathias Thoenig, and Thierry Verdier. 2014. "Product-Based Cultural Change: Is the Village Global?" *Journal of International Economics* 92: 212–30.
- Payutto, Prayudh A. (1998). *Buddhist Economics: A Middle Way for the Market*. Bangkok: Buddhadamma Foundation.
- Polanyi, Karl. 1944. *The great Transformation: The Political and Economic Origins of Our Time*. Boston, MA: Beacon Press.
- Pratt, Andy C. (2009). "Urban Regeneration: From the Arts 'Feel Good' Factor to the Cultural Economy: A Case Study of Hoxton, London." *Urban Studies* 46: 1041–61.
- Rao, Vijayendra, and Michael Walton (eds.). 2004. *Culture and Public Action*. Stanford, CA: Stanford University Press.
- Robbins, Lionel. 1932. *An Essay on the Nature and Significance of Economic Science*. London: Macmillan.
- Samuelson, Paul. 1947. *Foundations of Economic Analysis*. Cambridge, MA: Harvard University Press.
- Sandel, Michael J. 2012. *What Money Can't Buy: The Moral Limits of Markets*. New York: Farrar, Straus and Giroux.
- Scott, Allan J. 2004. "Cultural-Products Industries and Urban Economic Development: Prospect for Market Contestation in Global Context." *Urban Affairs Review* 39: 461–90.
- Scruton, Roger. 2010. *Liberty and Civilization: The Western Heritage*. New York: Encounter Books.
- Sen, Amartya. 2004. "How Culture Matters." In Vijayendra Rao and Michael Walton (eds.), *Culture and Public Action* (Stanford, CA: Stanford University Press), pp. 37–58.
- Simmel, George. 1900. *The Philosophy of Money*. Boston: Routledge & Kegan Paul (1978 edition).
- Smith, Adam. 1759. *The Theory of Moral Sentiments*. London: A. Millar.
- Storr, Virgil H. 2003. *Understanding the Culture of Markets*. New York: Routledge.
- Taylor, Charles. 1991. *The Malaise of Modernity*. Toronto: House of Anansi.
- Throsby, David. 2001. *Economics and Culture*. Cambridge: Cambridge University Press.
- Tinbergen, Jan. 1975. *On the Theory of Economic Policy*. Amsterdam: North-Holland.
- Titmus, Richard M. 1970. *The Gift Relationship: From Human Blood to Social Policy*. New York: New Press.
- Towse, Ruth. 2013. *A Textbook of Cultural Economics*. Cambridge: Cambridge University Press.
- Veblen, Thorstein. 1899. *The Theory of the Leisure Class*. New York: Macmillan.
- Visser, Hans. 2013. *Islamic Finance: Principles and Practice*. Cheltenham, UK: Edward Elgar.
- Walzer, Michael. 1983. *Spheres of Justice: A Defense of Pluralism and Equality*. Oxford: Basil Blackwell.
- Weber, Max. 1905. *The Protestant Ethic and the Spirit of Capitalism*. 3rd Roxbury edn. Stephen Kalberg (trans. and ed.). New York: Routledge (2012 edition).

- Webre, Yaakov. 1996. "Corporate Cultural Fit and Performance in Mergers and Acquisitions." *Human Relations* 49: 1181–202.
- Woodmansee, Martha, and Mark Osteen. 1999. *The New Economic Criticism: Studies at the Intersection of Literature and Economics*. London: Routledge.
- Zelizer, Viviana. A. 2005. *The Social Meaning of Money: Pin Money, Paychecks, Poor Relief, and Other Currencies*. Princeton, NJ: Princeton University Press.
- Zelizer, Viviana A. 2010. *Economic Lives: How Culture Shapes the Economy*. Princeton, NJ: Princeton University Press.
- Zuidhof, Peter W. 2012. *Imagining Markets: The Discursive Politics of Neoliberalism*. Rotterdam: Erasmus University.

PART II

APPLICATIONS

A. COMMERCE AND MARKETS

CHAPTER 14

HUMANE MARKETS

The Classical Tradition of Political Economy

JAMES R. OTTESON

14.1 INTRODUCTION

A common criticism of markets is that they are amoral, if not immoral. Skidelsky and Skidelsky (2012), for example, quote Lionel Robbins's famous definition of economics as "the science that studies human behaviour as a relationship between ends and scarce means which have alternative uses" (1932: 16) and state that "Robbins's definition both puts scarcity at the center of economics and brackets out judgments of value" (2012: 12–13). Many other commentators, from Marx to the present day, make similar claims.¹ Whatever the merit of these criticisms as they relate to contemporary economics, the central political economists of the eighteenth century took a different view.

In this chapter I argue that the core of the "political economy" that arose in the eighteenth century, which marked the start of the discipline we today know as economics, envisioned the exchanges that take place in commercial society as neither amoral nor immoral but indeed deeply humane. The claim of these early political economists was that business transactions in markets fulfilled principally two separate but related moral mandates: they lead to increasing prosperity, which addresses their primary "economic" concern of raising the estates of the poor; and they model proper relations among people, which addresses their primary "moral" concern of granting a respect to all, including the least among us. In this way, their argument attempted to capture a vision of human dignity within political-economic institutions that enabled people to improve their stations. Their arguments thus did not bracket out judgments of value: they integrated judgments of value into their foundations and built their political economy on that basis.

¹ See Marx (1844a), Ubel (2009), Satz (2010), and Sandel (2012).

To make that case, I review some core elements of the arguments from the two greatest political economists of the eighteenth century, David Hume and Adam Smith. They saw their projects in part as standing in opposition to Bernard Mandeville's claims about the selfishness and hypocrisy in commercial societies, a criticism that has animated numerous contemporary commentators as well. Thus I begin by sketching Mandeville's claims before turning to exploring Hume's and Smith's alternative vision of the commercial society. I then raise several worries about, and objections to, the commercial society they recommended, drawing on Mandeville's criticism and extending it to Rousseau, Marx, and a handful of contemporary critics. I close by suggesting that the classical political economists—primarily Hume and Smith—saw themselves as offering an attractive, even inspiring, vision of commerce and of the society that allows and encourages it.

14.2 HUMANE POLITICAL ECONOMY

Before the classical period of political economy had even started, Bernard Mandeville had already articulated a conception of commercial society that would serve as the basis of one of its most enduring criticisms. In his *The Fable of the Bees: Or, Private Vices, Publick Benefits*, Mandeville suggested that a commercial society rests on a deep hypocrisy: "All trades and Places knew some Cheat, / No Calling was without Deceit" (1723: 20).² Mandeville's claim was that all trades, regardless of their public claims and apologies, operated in fact on the basis of narrow self-interest. Laborers, artificers, merchants, and professionals were all interested first and foremost in satisfying their own ends, in achieving gain for themselves. In a commercial society, however, they could not achieve their goals by means of enslavement, theft, or fraud, which meant that the only option left available to them was to convince others to trade or exchange with them voluntarily; that, in turn, meant that they would have to flatter the sensibilities of potential clients or customers by pretending to care about their interests, rather than only about their own. But this was a false pretense. According to Mandeville, even vocations that were then, and still today, considered to be fundamentally other-regarding—including everything from judges to physicians to priests and clergy—were still, regardless of their pretenses to the contrary, only concerned for themselves. Judges and lawyers might claim they are interested in justice, but if there were no injustice, they would have no work; physicians, for their part, claim to be interested in the health of their patients, but if there were no disease or sickness, they too would have no work; and what would priests or clergy do if people did not in fact continue to be sinful?³ Yet because people wish to be flattered, and

² This work first appeared as the poem "The Grumbling Hive: Or, Knaves Turn'd Honest" in 1705; Mandeville expanded it and reissued it under the title *The Fable of the Bees* in 1714. I quote from the 1723 edition, which was reissued by Oxford in 1924 and reprinted by Liberty Fund in 1988.

³ Perhaps it is noteworthy that Mandeville himself was a physician.

prefer not to think that others do not in fact care about them, this gives rise to a systematic duplicity: everyone claims to care about others so that they can get what they want, but in fact no one cares about anyone other than himself. As Mandeville claims, in a commercial society, “Thus every Part was full of Vice” (1723: 24).

The implications of Mandeville’s argument went beyond merely the claim that everyone is fundamentally self-interested, however. He argued that it was precisely this vice of selfishness that enabled the economic prosperity—the “publick benefit”—that we all enjoy. In Mandeville’s poem, Jove grew angry at the hypocrisy he saw, and so he made the fateful decision to grant people’s publicly expressed wish to no longer have the vices of injustice and sin. When Jove does so, there suddenly ensues a virtual economic collapse. Because there was now no more sin, there was now no longer any need for judges, physicians, and clergy. Similarly, because there was now no envy, greed, ambition, or avarice, many of the publicly decried “luxuries” went away too, leading to the death of industries like architecture and large-building manufacturing, ship-building, the making of fine foods and wines and clothing, seafaring trade, and so on. Because everyone was now honest, content, and virtuous, “All Arts and Crafts neglected lie; / Content, the Bane of Industry, / Makes ‘em admire their homely Store, / and neither seek nor covet more” (1723: 34–5).

Mandeville’s argument is hence not only that people are in fact vicious insofar as they are narrowly self-interested yet lie about it, but that it is good that they are—for otherwise we would not have the goods and services, and the prospering society, that we all want and from which we all benefit. The apparent moral to draw is that economic prosperity is possible only if (1) people embrace their fundamentally self-interested motivations and (2) continue to pretend otherwise. Later critics of commercial society would repeat claims like these to establish the alleged amorality, even immorality, of markets. Before we look at the use to which later critics put those claims, however, let us look at the very different conception of commercial society we get from Hume and Smith.

14.2.1 Hume’s Humane Political Economy

In his essay “Of Refinement in the Arts,” Hume takes up the pressing question of “luxury.” In his day, luxury was considered a vice, a “great refinement in the gratification of the senses” that often indicated licentiousness, greed, superficiality, or undue concern with trivial matters (1754a: 268). Hume argues, however, that luxury in fact “may be considered either as innocent or blamable,” depending on the circumstances; indeed, he argues that “the ages of refinement are both the happiest and most virtuous” and that “wherever luxury ceases to be innocent, it all ceases to become beneficial”—implying that “innocent” luxury is in fact beneficial (269). He goes so far as to claim that “*industry, knowledge, and humanity*, are linked together by an indissoluble chain, and are found, from experience as well as reason, to be peculiar to the more polished, and, what are commonly denominated, the more luxurious ages” (271, italics in the original). Hume’s claim is that the desire for luxuries leads to the development of both mechanical and fine

arts to supply these desires. This leads to increasing “industry” and “knowledge.” Even more important, in order to create these luxuries people must cooperate with one another—including with people from different places, with differing nationalities, languages, and cultures—which tempers their enthusiasms and prejudices and makes them “more sociable” (*ibid.*). This leads to increasing “humanity.”

Hume argues that the benefits from industry, knowledge, and humanity accrue not only to individuals, however, but also to the public. They “serve to the ornament and pleasure of life” generally, and the increasing prosperity that they portend generates “a kind of *storehouse* of labour, which, in the exigencies of state, may be turned to the public service” (272; *italics in the original*). This, according to Hume, is in fact a multifold benefit: (1) individuals become more “happy and prosperous,” as they see more of their needs and desires met; (2) as knowledge and luxuries increase, individuals become refined and elevated in their thoughts, manners, and sentiments; and (3) as overall prosperity increases, states have more reserves on which to draw to engage in beneficial large-scale projects like infrastructure, education, and even, when necessary, war. Perhaps most important for Hume, the benefits accrue particularly for the lowest in society: “But where luxury nourishes commerce and industry, the peasants, by a proper cultivation of the land, become rich and independent; while the tradesmen and merchants acquire a share of the property, and draw authority and consideration to that middling rank of men, who are the best and firmest basis of public liberty” (277).

These are moral considerations for Hume. Bettering the conditions of all of us, including especially the poor; increasing the knowledge and skills in society; softening prejudices and making people “more sociable”; and increasing productivity and thus enabling general prosperity are all moral values that Hume believes are worth championing. But Hume finds still other benefits in a commercial society. In another essay, he writes: “Nothing is more usual, among states which have made some advances in commerce, than to look on the progress of their neighbors with a suspicious eye, to consider all trading states as their rivals, and to suppose that it is impossible for any of them to flourish, but at their expense” (1754b: 328). This is an example of the “zero-sum” fallacy, or the (false) belief that mutually voluntary exchanges benefit one party only at the expense of the other. Hume argues instead that mutually voluntary exchanges are mutually beneficial, or “positive-sum”: “In opposition to this narrow and malignant opinion, I will venture to assert, that the encrease of riches and commerce in any one nation, instead of hurting, commonly promotes the riches and commerce of all its neighbours; and that a state can scarcely carry its trade and industry very far, where all the surrounding states are buried in ignorance, sloth, and barbarism” (*ibid.*).

Thus the benefits Hume sees from commerce reach from the individual to the nation to the international community, and Hume’s argument was intended in no small part to encourage his readers to view others’ success—including the success of other countries—not “with a suspicious eye” that arises from “jealousy,” but, rather, with a generosity of soul that understands that we “cannot be hurt by the greatest prosperity of [our] neighbors” (*ibid.*). The softening of prejudices to which Hume argued commerce can lead would also, he suggested, lead to a tempering of nationalism. It turns out that

“Nature, by giving a diversity of geniuses, climates, and souls, to different nations, has secured their mutual intercourse and commerce, as long as they all remain industrious and civilized” (329). So the differences among peoples are not, or should not be, causes for concern or alarm, but instead causes of optimism: our differences complement one another, and thus we can benefit from one another. “They consume the produce of my industry, and afford me the produce of theirs in return” (ibid.). Hume’s conclusion: “I shall therefore venture to acknowledge, that, not only as a man, but as a BRITISH subject, I pray for the flourishing commerce of GERMANY, SPAIN, ITALY, and even FRANCE itself” (331; *sic*). Yes, even France: Hume thought that commerce could bring even longstanding enemies to see each other not as threats or rivals, but rather as opportunities.

Although Hume does not in these essays articulate the specific institutional requirements that would enable the peaceful mutual flourishing to which he thinks commerce can lead, his call for free trade and commerce would seem to require the reduction or elimination of legally enforced trade barriers, protections from competition, and tariffs, as well as for the protection of private property rights and the enforcement of voluntary agreements and contracts.⁴ But all of the benefits Hume identifies as potentially arising from commerce are motivated by exactly the kind of moral concerns that Skidelsky and Skidelsky (2012) believe the economist “brackets out”: concern for the wellbeing of the individual, of other individuals, of the nation, and indeed of the world. For Hume, this is not driven by the narrow sense of self-interest Mandeville described, but is, instead, driven by an expanded self-interest that finds benefit, and even moral worth, in serving one’s own aims only by serving those of others.

14.2.2 Smith’s Humane Political Economy

The reader may have detected a nascent “invisible hand” argument in Hume’s account: by pursuing their own direct interests within the context of a properly configured commercial society, individuals engage in activity that also conduces both to indirect benefits to themselves and to the benefit of society more generally. The direct benefit to the individual is the obtainment of the particular good or service she desires; the indirect benefit is the expanding of horizons and circles of sympathy, and the softening of prejudices and of nationalisms, to which commerce can lead an individual, even unwittingly. The benefit to society is the increased stock of goods and services—that is, increasing prosperity—of the society, as well as an easing of interest in foreign antagonisms and war.

Adam Smith would go on to make a similar argument, but he expands on the way he believes commerce can lead to better relations among people. He also specifies more concretely what the institutional requirements are for such mutually beneficial outcomes. Consider these in turn.

⁴ See Hume (1740), however, where he specifies “the three fundamental laws of nature, *that of the stability of possession, of its transference by consent, and of the performance of promises*” (337, italics in the original).

In what is perhaps the second-most-famous passage from his 1776 book *An Inquiry into the Nature and Causes of the Wealth of Nations* (WN), Smith writes: “It is not from the benevolence of the butcher, the brewer, or the baker, that we expect our dinner, but from their regard to their own interest. We address ourselves, not to their humanity but to their self-love, and never talk to them of our own necessities but of their advantages” (WN: 26–7). As famous as this passage is, its meaning is contested. There is the question, first, of whether Smith is making a descriptive or a normative claim: is he merely describing how people in fact behave in markets, or is he making a recommendation about how they ought to behave? A second question, which is perhaps more relevant to our discussion, is what exactly Smith is identifying as the proper motivation in market transactions. One possibility is selfishness. Smith does, after all, use the phrase “self-love,” and he specifically disclaims addressing “humanity.” One interpretation, then, is that Smith is offering, barely twenty pages into *The Wealth of Nations*, not only a recognition of the role selfishness might play in market exchanges but might actually be advocating selfishness in them. If Smith is claiming, in effect, that one must check one’s concern for others at the market door, then one might well conclude: so much the worse for the market.

Here is a different interpretation. What Smith saw himself as advocating was, not selfishness, but respect. Just before the butcher-brewer-baker passage, Smith says that human beings, unlike other animals, must rely on others to acquire what they want. “Nobody ever saw a dog make a fair and deliberate exchange of one bone for another with another dog” (WN: 26), because they do not need to: dogs, like most other animals, can procure most of what they need all on their own. “But man has almost constant occasion for the help of his brethren” (ibid.), both because our needs and wants are more complex than those of other animals and because human beings do not have the equipment—fur, claws, wings, and so on—that nature provided other animals to enable them to satisfy their needs. What do human beings have to compensate for their relative physical weaknesses? They have “the faculties of reason and speech” (WN: 25), which enables them to discover and construct plans for cooperating with one another in ways that makes all parties better off. This is why “man has almost constant occasion for the help of his brethren.” But to this Smith adds: “and it is in vain for him to expect it from their benevolence only. He will be more likely to prevail if he can interest their self-love in his favour, and shew them that it is for their own advantage to do for him what he requires of them” (WN: 26). How can one do this? Smith is suggesting that the way to do this is by offering to the other person something that that other person values: “Give me that which I want, and you shall have this which you want, is the meaning of every such offer” (ibid.). Because of our peculiar liabilities and limitations, human beings need the help of others; and it is by making mutually advantageous offers “that we obtain from one another the far greater part of those good offices which we stand in need of” (ibid.).

The assumption Smith makes in this argument is indeed that we are driven by self-interest.⁵ But the exchanges he describes in this passage are constrained by two

⁵ Elsewhere in WN Smith speaks of individuals’ “universal, continual, and uninterrupted effort to better their condition,” which he says is “a desire which, though generally calm and dispassionate, comes with us from the womb, and never leaves us till we go into the grave” (WN: 345, 341).

other claims he makes: (1) the strong and instinctive desire each of us has for what Smith calls, in his 1759 book *The Theory of Moral Sentiments* (TMS), “mutual sympathy of sentiments” (9–16) and (2) the requirement of living in a “well-governed society” (WN: 22). In TMS, Smith argues that “nothing pleases us more than to observe in other men a fellow-feeling with all the emotions of our own breast” (TMS: 13). When we observe that our sentiments are in “sympathy” (or “harmony,” “concord,” or “correspondence,” other terms Smith uses) with others’ sentiments, it gives us a great pleasure, just as when we observe that our sentiments do not share a sympathy with those of others it gives us displeasure. An example Smith uses to illustrate the principle: “A man is mortified when, after having endeavoured to divert the company, he looks round and sees that nobody laughs at his jests but himself. On the contrary, the mirth of the company is highly agreeable to him, and he regards this correspondence of their sentiments with his own as the greatest applause” (TMS: 14). Smith believes that it is an empirical fact of human psychology that we take great pleasure in having our actions, judgments, and sentiments not only approved by others but also, even further, in being worthy of that approval. This leads us to be genuinely concerned about others: “Man naturally desires, not only to be loved, but to be lovely; or to be that thing which is the natural and proper object of love” (TMS: 113). So we do in fact, Smith claims, care about what others think of us, we do in fact strive to be worthy of their approval, and thus we do in fact take an interest in others.⁶ The first sentence of TMS is: “How selfish soever man may be supposed, there are evidently some principles in his nature, which interest him in the fortune of others, and render their happiness necessary to him, though he derives nothing from it except the pleasure of seeing it” (TMS: 9).

The second assumption behind Smith’s butcher-brewer-baker passage is that such exchanges take place in a “well-governed society” (WN: 22). What Smith means by this is also first articulated in TMS. It is a society that enforces Smith’s conception of “justice,” which for him comprises three “laws”: (1) “the laws which guard the life and person of our neighbour; [2] the next are those which guard his property and possessions; and [3] the next are those which guard what are called his personal rights, or what is due to him from the promises of others” (TMS: 84). A well-governed society, then, is one that protects “justice”—that is, in which people’s persons, property, and promises are protected, and in which there is what Smith calls a “regular administration of justice” (WN: 565) to address infractions of “justice.” In such a society, one cannot get what one wants from others by slavery, imperialism, colonialism, theft, fraud, or any other extractive means, because all of these are debarred by the requirements of “justice.” But the nature of humanity’s limitations requires that we must rely on others to achieve what we want. What route, then, is left to get what one wants from others? Smith’s answer: by making them an offer, which they are free to decline if they so choose. Smithian justice ensures what we might call an “opt-out option,” or everyone’s right to say “no, thank you” to any offer or request. To deny this right—that is, to mandate, force, or coerce a transaction with an unwilling partner—would constitute an infraction of justice. Thus the only

⁶ See Otteson (2013).

allowable transactions in a Smithian “well-governed society” are those that are mutually voluntary, and hence mutually beneficial according to each party’s individual conception of benefit and schedule of value. Smithian justice rules out extraction and therefore allows and endorses instead only voluntary cooperation.

How do these two constraints—the “desire for mutual sympathy of sentiments” and a “well-governed society”—relate to the butcher-brewer-baker passage? Smith’s argument is that in such transactions, we are driven, both by our own desires and by the constraints of our public institutions, to meet one another as peers, as moral equals, and to make offers to one another that either of us is free to decline. Because each of us has an “opt-out option” that is protected by our society’s commitment to Smithian justice, this disciplines us from any notion we might otherwise have had about trying to coerce, steal from, or defraud one another. And because each of us desires mutual sympathy of sentiments, we desire to conduct ourselves in ways that others will approve of. When we seek our meat from butchers, our ale from brewers, and our bread from bakers, then, we make them offers that recognize that they are our equals, that they have legitimate interests and obligations of their own, and that our interests and obligations do not trump theirs. Our desire for their meat, ale, and bread—which they had to make with their own labor and time and resources—therefore may not trump their right to decide on their own what to do. In these circumstances, we can get their meat, ale, or bread only by treating them the way they want to be treated. We will thus have to offer them something they might want, and they, for their part, will have to do the same to us—otherwise each of us will go elsewhere. In other words, we have to treat each other with respect, and not presume that either of us is more important or more worthy or more deserving than the other. And we do this by expecting our dinner “not from the benevolence of the butcher, the brewer, or the baker,” but, rather, by taking due and proper “regard to their own interest.”

For Smith, then, the act of making a person an offer is a recognition of the equal moral agency of others. It reflects the dignity that each of us has, and it is a paradigmatic example of proper moral relations among people. The mutually voluntary and thus mutually beneficial transactions that are the cornerstone of a Smithian market economy are, then, not only the key to increasing general prosperity, but for him they are also the instantiation of properly moral human relations.

The most famous passage in Smith’s WN, the “invisible hand” passage, is often taken to represent merely economic gain: by seeking to satisfy my own ends, I am “led by an invisible hand” to promote an end which was no part of my intention, namely, “to promote the publick interest” (WN: 456). The standard interpretation of this passage is that Smith is claiming that individual profit-seeking will lead to an overall increase in the wealth or prosperity of society. That surely is his claim, but it is only part of it. Putting the “invisible hand” passage together with the butcher-brewer-baker passage, we can see that Smith is also making a moral claim of equality: we are equal moral agents, each recognizing the other as a moral peer who deserves just as much respect for her choices as we deserve for our own. Moreover, the unintended benefit to others from our cooperative efforts to secure our own ends is itself a moral good—because those others matter too.

Here, then, we see several of the many agreements between Smith and Hume. Hume writes that in market transactions, others “consume the produce of my industry, and afford me the produce of theirs in return” (1754b: 329); Smith writes, “Give me that which I want, and you shall have this which you want, is the meaning of every such offer” (WN: 26). Hume writes, “Nature, by giving a diversity of geniuses, climates, and soils, to different nations, has secured their mutual intercourse and commerce” (1754b: 329); “Among men,” Smith writes, “the most dissimilar geniuses are of use to one another; the different produces of their respective talents, by the general disposition to truck, barter, and exchange, being brought, as it were into a common stock, where every man may purchase whatever part of the produce of other men’s talents he has occasion for” (WN: 30). Hume claims that “where an open communication is preserved among nations, it is impossible but the domestic industry of every one must receive an increase from the improvements of others” (1754b: 328); and Smith claims that “Every workman [...] supplies [others] abundantly with what they have occasion for, and a general plenty diffuses itself through all the different ranks of the society” (WN: 22). Thus: reciprocity, mutual respect, and improvement in the lives of everyone in one’s society and of those in other societies. For Hume and Smith, then, the political-economic recommendations they make about private property, trade, and production are not only not amoral (let alone immoral), but are on the contrary driven by what they see as deep concerns about proper moral relations among people and the moral mandate to better the condition of others, especially the least among us.

14.3 CRITICISMS OF HUMEAN AND SMITHIAN POLITICAL ECONOMY

Neither Hume nor Smith was a Pollyanna, however: both of them were conscious of the many ways that human beings behave corruptly, viciously, or ignorantly. Hume, for example, went so far as to write: “It is, therefore, a just *political* maxim, *that every man must be supposed a knave*: Though at the same time, it appears somewhat strange, that a maxim should be true in *politics*, which is false in *fact*” (1741a: 42, italics in the original). And Smith, for his part, spent a great deal of time discussing the knavery of not only princes, statesmen, and legislators, but of businesspeople as well: “People of the same trade seldom meet together, even for merriment and diversion, but the conversation ends in a conspiracy against the publick, or in some contrivance to raise prices” (WN: 145).⁷ Nevertheless, both Hume and Smith believed that a properly ordered commercial society would lead not only to the “economic” benefit of increasing overall wealth and prosperity

⁷ See, for example, Smith’s discussion in TMS of the statesman who is a “man of system” (TMS: 233–4) and his discussion in WN of the “dangerous” “folly and presumption” of statesmen (WN: 456).

but also to the “moral” benefits of increasing virtue, decreasing vice, and increasing respect for moral equality.

Not everyone has been convinced, however, particularly of the latter claims about morality. Indeed, the criticisms of, worries about, and objections to markets and commercial society are legion—far too many to address in one essay. But let us raise and address three that have inspired both historical and contemporary concerns: (1) that markets are dehumanizing; (2) that markets depend on moral vice; and (3) that markets lead to an objectionable inequality.

14.3.1 Dehumanizing Markets

For this objection, we can draw on a contemporary of Hume and Smith, Jean-Jacques Rousseau, as well as a trenchant nineteenth-century critic of their vision of political economy, Karl Marx.

Rousseau worried that commercial society would lead to an interdependence that he found loathsome. In his 1755 *Discourse on the Origin of Inequality*, Rousseau imagined—and even seemed to idolize—his vision of man in his natural state, which was one of an almost carefree independence: “by considering him, in a word, such as he must have issued from the hands of Nature, I see an animal less strong than some, less agile than others, but, all things considered, the most advantageously organized of all: I see him sating his hunger beneath an oak, slaking his thirst at the first Stream, finding his bed at the foot of the same tree that supplied his meal, and with that his needs are satisfied” (134). For this “Savage man,” as Rousseau calls him, his “modest needs are so ready to hand” (143); indeed, “males and females united fortuitously, according to chance encounters, opportunity, and desire, without speech being an especially necessary interpreter of what they had to tell one another; they parted just as readily” (145). In Rousseau’s vision, “Savage man” needs nothing beyond what his own abilities can procure for him, and this grants him both freedom and independence: “without any need of others of his kind and without any desire to harm them, perhaps even without ever recognizing any one of them individually, subject to few passions and self-sufficient, Savage man had only the sentiments and the enlightenment suited to this state, that he sensed only his true needs” (157).

Yet as Rousseau puts it in his 1762 *Of the Social Contract*, “Man is born free, and everywhere he is in chains” (41). The chains come principally from commercial society, which takes man out of his free and independent natural state and makes him dependent on others to satisfy his needs and wants. “As he becomes sociable and a Slave, he becomes weak, timorous, groveling, and his soft and effeminate way of life completes the enervation of both his strength and courage” (1755: 138–9). In the *Discourse* Rousseau lays a great proportion of humanity’s ills at the feet of commercial society—everything from disease to war. The crux of his criticism is that commercial society dehumanizes us by making us dependent on one another, thus rendering us not only enervated and weak but, as he puts it, even “slaves” to one another.

The claim that commercial society leads to interdependence, if not dependence, is not only recognized by Hume and Smith, however—it is celebrated. In WN, for example, Smith describes the “woollen coat” that “the day-labourer” wears. When one considers all the people and hands that cooperated to make this coat and get it to the day-laborer, Smith says it “exceeds all computation” (WN: 22). Smith does not think this is an exaggeration. He writes:

The shepherd, the sorter of the wool, the wool-comber or carder, the dyer, the scribbler, the spinner, the weaver, the fuller, the dresser, with many others, must all join their different arts in order to complete even this homely production. How many merchants and carriers, besides, must have been employed in transporting the materials from some of those workmen to others who often live in a very distant part of the country! How much commerce and navigation in particular, how many ship-builders, sailors, sail-makers, rope-makers, must have been employed in order to bring together the different drugs made use of by the dyer, which often come from the remotest corners of the world! What a variety of labour too is necessary in order to produce the tools of the meanest of those workmen! (WN: 22–3)

Smith continues at some length in tracing out the links in this surprisingly complex and far-flung chain of interdependence. For Smith, this is a great triumph of cooperation, traversing and breaking down barriers of nationality, religion, and culture. Yet to Rousseau, interdependence is still dependence, and hence he would prefer to forgo the coat—and the other goods and services commercial society provides—rather than to suffer what he believes is the dehumanizing reliance on others that it entails.

Karl Marx raises a related criticism of dehumanization, but he offers a different conception of it and a different argument for it. More familiar Marxian criticisms of markets relate to the “alienation” and “commodification” to which he claims markets lead. A somewhat less familiar criticism, though in some ways more damning, comes from his early commentaries on James Mill. There Marx writes:

When I produce *more* of an object than I myself can directly use, my *surplus* production is cunningly *calculated* for your need. It is only in *appearance* that I produce a surplus of this object. In reality I produce a *different* object, the object of your production, which I intend to exchange against this surplus, an exchange which in my mind I have already completed. The *social* relation in which I stand to you, my labour for your need, is therefore also a mere *semblance*, and our complementing each other is likewise a mere *semblance*, the basis of which is mutual plundering. The intention of *plundering*, of *deception*, is necessarily present in the background, for since our exchange is a selfish one, on your side as on mine, and since the selfishness of each seeks to get the better of that of the other, we necessarily seek to deceive each other. (1844a, italics in the original)

There are in fact several criticisms of markets in this passage, but let us focus on just one: Marx’s claim that in market exchanges the “plundering” and “deception” in which we engage indicates a violation of the Kantian categorical imperative to treat others as ends

in themselves and never merely as means. Marx's argument is that the negotiation that takes place in markets—what Smith calls “the higgling and bargaining of the market” (WN: 49)—proceeds on the basis of lying. You say to the car salesman, for example, “I will not pay more than \$20,000 for that car.” But that is a lie. The salesman replies, “We can't take less than \$25,000 for the car.” That too is a lie. And so it goes, throughout the negotiation process.

This is a recapitulation of Mandeville's claim of hypocrisy in commercial society. Even if an agreement is reached, and even if the agreement is mutually voluntary and mutually beneficial, still it is effectuated by mutual lying, and thus of each party selfishly treating the other as a mere instrument to be manipulated to achieve one's own goals. I want the car, and you want to sell it to me; thus, Marx argues, we are induced to manipulate one another through deception in our attempts to achieve our respective ends. The dehumanization that this constitutes is reflected in the extent to which Marx thinks markets induce us to see others not as full moral agents, not as persons with dignity we are morally bound to respect, but rather as tools like any other to be used for our purposes. Perhaps we are more complicated tools than most others—which is why we need to “calculate” and “deceive”—but we treat each other as tools nonetheless.

14.3.2 The Vice of Markets

Marx's claim that markets dehumanize us by inducing us to treat one another as mere tools implicates the extent to which he also believes that market transactions proceed on the basis of vice. Lying is morally wrong, after all. If we are in the grips of a capitalist ideology, we might attempt to rationalize our behavior; perhaps we say “everybody does it.” But that does not mean it is, or should be, acceptable. Perhaps bluffing in poker is acceptable, where either the stakes are low or we can be sure that in fact everyone knows that bluffing is part of the game (or both); but our standards for proper behavior generally, including our public institutions, specifically forbid, or should forbid, systematic lying. Marx's criticism of markets here goes beyond just decrying particular instances of lying, however. His suggestion is that because we can get rewarded for lying in markets, and the better we are at lying the bigger the reward will be, lying becomes habitual. It becomes an endemic part of our personalities. But once lying becomes a part of who we are, it negatively affects other aspects of our personalities; in this way, immoral behavior that is rewarded by markets can poison our souls, as well as our relationships with others. This can be another instance of the “alienation” to which Marx argues markets leads: our relations with others become so corrupted and corroded that we can no longer experience the deep community with others—we cannot partake in our “species being”—that enables a properly flourishing life.⁸

The language of corruption comes up in contemporary critics of Humean/Smithian markets as well. Michael Sandel (2012), for example, worries about “the corrosive tendency

⁸ See Marx (1844b: 58–68).

of markets,” claiming that “markets don’t only allocate goods; they also express and promote certain attitudes toward the goods being exchanged” (9). Those attitudes are a function of “the logic of buying and selling,” and they “don’t ask whether some ways of valuing goods are higher, or worthier, than others” (6, 14). Sandel continues that the “corruption objection” to markets “points to the degrading effect of market valuation and exchange on certain goods and practices. [...] Certain moral and civic goods are diminished or corrupted if bought and sold” (111). Sandel gives numerous examples of things that are “corrupted if bought and sold,” including everything from paying to hunt walrus to charging for admission to Yosemite National Park to buying or selling semen or eggs to scalping tickets to see the Pope: “Treating religious rituals, or natural wonders, as marketable commodities is a failure of respect. Turning sacred goods into instruments of profit values them in the wrong way” (37).

Sandel does not give much substantive guidance about what would constitute valuing such things in the right way (or what would be an effective mechanism for adjudicating the inevitable disagreements about what the “right” value for a good is), but his worries about corruption are shared by many others. Debra Satz (2010), for example, also worries about the corrosive effects of what she calls “noxious markets” in things like women’s reproductive labor and sexual labor. According to Satz, a “noxious market” is one that leads to harmful outcomes for individuals, that leads to harmful outcomes for society, that is characterized by weak agency or asymmetric knowledge, or that reflects vulnerabilities of some parties to exchanges (94–100). Although Satz argues that surrogacy, the selling of eggs, and prostitution are prime examples of “noxious markets”—she writes that “women’s reproductive labor is intrinsically not a commodity” (115)—her criteria for determining when a market is “noxious” are quite broad and would include markets for many other goods and services. Financial services and medical care, for example, would often trigger the knowledge-asymmetry criterion, and the criterion of possibly leading to harmful outcomes for individuals or for society might jeopardize everything from the sale of guns and tobacco to motorcycles and cars to alcohol and (many) foods to even, possibly, some kinds of music, movies, literature, and social media. All of these things have been argued by various people to produce harmful outcomes—even “extremely harmful outcomes” (94)—which would seem to imply, on Satz’s argument, the curtailing of in fact a great deal of what gets bought and sold in markets.

Other recent critics worry about the corroding effects of pornography, of paying students to read books, of paying people to stand in line for oneself, of paying for first-class airplane seats, of paying for human-hair wigs, of paying for kidneys or blood, of various kinds of sexual services, and many other things—some too graphic to be described here—that one can find available for sale in the world today. Most such critics call for government regulation or for outright prohibition of the buying or selling of such things, on grounds that they are disgusting, revolting, disrespectful, profane, bizarre, perverse, or unworthy.⁹

⁹ For these and other examples of condemned markets, as well as these terms used to disapprove of them, see Radin (2001), Schor (2004), Barber (2008), Ubel (2009), Satz (2010), Sandel (2012), Skidelsky and Skidelsky (2012), and Conly (2013). For a discussion of many of these examples, see Brennan and Jaworski (2016).

Not all the examples are salacious or graphic, however. Medical doctor Peter Ubel (2009), for example, argues that because, as we now know, people often make irrational or biased decisions, we should restrict many of their choices for behaviors that could lead to obesity, including drinking sugary sodas, eating doughnuts and other pastries, and playing video games.¹⁰ Philosopher Sarah Conly (2013) goes even further, calling for complete governmental “coercive paternalism,” to relieve individuals of the burden of having to make their own decisions and—apparently assuming flawless, or nearly flawless, governmental decision-making—the costs resulting from bad decisions. Even more recently, Conly (2015) calls on the government to intervene in the decision of whether to have children or not, arguing that the government can make better decisions in this matter than individuals can or do.

Though their arguments are diverse, what these critics share is a concern that unregulated markets—markets, that is, in which individuals or private parties make mutually voluntary transactions not subject to uninvited third-party interposition—will proliferate in exchanges that the critics disapprove. More than that, however, they agree that these disapproved exchanges will have deleterious effects on the individuals involved or on the culture more generally. They will lead to a coarsening of manners, to the disrespecting of proper value, and to various kinds of vice, quite apart from any potential misallocation of resources, negative externalities, or decreases in overall prosperity.¹¹

14.3.3 Inequality

Scholars worried about inequality employ many different conceptions of equality. Some worry about inequality of wealth or income, some about inequality of capacity, and some about inequality of outcomes.¹² Does commercial activity in markets lead to at least some kinds of inequality? Yes. Because people are different, because people’s purposes, values, goals, and opportunities are different, because people’s skills and abilities are different, and because of sheer luck, people’s lives will take different, and unique, trajectories. That means they will have differing outcomes in life, including differing levels of material wealth.

Perhaps such material inequality is inherently unjust.¹³ It would not be unjust on a Smithian conception of justice, however, which, as we have seen, entails only protection from injury—specifically, injury in our persons, our property, or our voluntary promises.

¹⁰ See also Thaler and Sunstein (2009) and Thaler (2015).

¹¹ Though we cannot pursue it here, I note that there is now substantial empirical evidence that markets increase prosperity generally, including especially for the poor. Moreover, there is also evidence suggesting that markets actually lead people to behave with greater fairness, to cheat less, to be inclined to trust others more, to be more altruistic, and to be in fact less inclined to lie. For the former, see, for example, Hall and Lawson (2014) and McCloskey (2016); for the latter, see Zak (2008, 2012), Gintis (2012), Hoffman and Morgan (2013), and Brennan and Jaworski (2016).

¹² See, respectively, Tomasi (2012) and Piketty (2014), Sen (1992) and Nussbaum (2011), and Fleurbaey (1995).

¹³ See Frankfurt (1998).

But some argue that we should expand the conception of justice, perhaps to include a prohibition of (suitably extreme) inequality. To illustrate the Smithian position, consider two people, A and B. A has more wealth than B—much more wealth. Let us say that A has 100 times as much wealth as B. Without knowing anything more about these two people, should we believe that there is something necessarily wrong with a society in which one person can have 100 times as much wealth as another? Let us further suppose that A did not steal from or defraud anyone to get her wealth; she engaged only in transactions that created value not only for herself but for others, according to those others' own judgment. Some, like Piketty (2014), would claim that there is something *prima facie* wrong with A having so much more than B.

Consider, however, that if A has this much more than B, it might mean that B's life is miserable. Perhaps having only 1/100th as much as A means that B lives a life of poverty and misery, below the standards of basic human decency. Perhaps B's options in life and opportunities for improvement are far fewer than those of A—which, it seems, would almost certainly have to be the case. Is B, then, owed something? If not from A directly, then perhaps from society?

I suggest that Hume and Smith would answer in the negative. To see why, suppose our A is Bill Gates, who, as of this writing, has a net worth of approximately \$90 billion; and our B is Michael Jordan, who, as of this writing, has a net worth of approximately \$900 million, or 1/100th of that of Gates. Would this great inequality between Gates and Jordan, and in particular Jordan's extreme relative poverty—relative, that is, to Gates, not relative to the rest of us—entail that he is entitled to some compensation? Obviously not. Jordan might be worth only a tiny fraction of what Gates is, but he is doing just fine on his own by any reasonable standard. Now, this is an extreme case, but it suggests the general point that the fact that one person has more wealth than another person does not, by itself, suffice to initiate compensatory action. To justify compensation or redistribution of wealth on the Humean/Smithian argument, we would need to show additionally that A got his wealth through unjust means (violation of person, property, or promise; through threats, force, or fraud) or that B has been prevented from increasing his wealth through similarly unjust means. But in such cases it is not the *inequality* that requires rectification but the *injustice*—and that is already addressed by Smithian justice.

As the history of human prosperity, as well as the results of numerous historical experiments in differing systems of political economy, seem to have shown, we seem indeed to face a choice between general prosperity and material equality. For most of human history, people were relatively equal in levels of wealth, but they were equally poor. During the only time in history that substantial numbers of human beings were able to rise out of poverty—since around 1800—what enabled this rise also led to material inequality.¹⁴ The lesson from human economic history seems to be that the only way we have ever discovered to enable substantial numbers of people to rise out of poverty is a set of political-economic and cultural institutions that also engender inequality.

¹⁴ See McCloskey (2006, 2010, 2016). See also Landes (1999) and Clark (2007).

Conversely, equality has historically correlated with low levels of prosperity. As Scheidel (2017) has recently argued, the central historical causes of reductions in inequality, or the great causes of equality—what Scheidel calls the “Four Horsemen” of leveling—are war, violent revolutions, state collapse, and catastrophic plagues. All of these are combatted by having a government that protects justice and by the increasing prosperity such a society fosters. This suggests that we seem to face this stark choice: we can combat poverty (by increasing general prosperity), or we can combat inequality (by extraction, coercion, or poverty). If you could solve only one social ill, either poverty or inequality, which would you choose? This is not an idle question, for it seems to be the actual dilemma facing humanity. We can address one or the other, but not both. Hume and Smith would argue that poverty is more deserving of our concern. Poverty is the cause of far greater misery, and what keeps people in poverty—Scheidel’s “Four Horsemen” listed above—are far more destructive of human flourishing than is inequality. This is especially true if the institutions enabling mass flourishing actually benefit everyone, including especially the least among us, even if it does not benefit everyone equally. And that has actually been the case in the over two centuries since Smith published his *Wealth of Nations* for the people lucky enough to have been members of those countries that adopted his prescriptions.

The Humean and Smithian predictions were that a society constituted with the right institutions and with members possessing the right characters will see increasing wealth. In fact, their argument seemed to suggest that there was no upward bound on the wealth—and genuine prosperity—that such a society could generate and enjoy. Even a lone country with these characteristics, surrounded by other countries engaging in extraction, cronyism, corruption, and so on would still soon outpace the others in prosperity. As both Hume and Smith argued, it would then become an example to those other countries, as well as a beacon welcoming trade with and commerce from other countries. In this way it would place competitive pressures on other countries to give up their extractive ways, on pain of losing people or resources or both, and thus to reform their institutions and attitudes to encourage cooperative and mutually beneficial interactions.

14.4 CAN MARKETS BE HUMANE?

There are many other critics of the Humean/Smithian commercial society, and of markets, business, and economics more generally—too many, in fact, to address here. And we cannot hope to respond to their worries and objections within the compass of one chapter. We will have to leave to the reader to imagine how Hume and Smith might respond, and whether their responses would be convincing. But let us raise one final objection to markets that has been raised by a number of recent critics, an objection to which a Humean/Smithian answer will give a good indication of how the latter’s

argument in favor of commercial society captures the sense in which they believe commercial society can actually conduce to a just and humane society.¹⁵

There is a position in political and moral philosophy known as *luck egalitarianism*. Although it has many variations, its core holds that luck plays a significant role in the relative levels of success people achieve in life.¹⁶ It further claims that, because luck by definition is unearned and thus undeserved, people who have experienced good luck have no more claim to be entitled to its benefits than people who have experienced bad luck should have to suffer its costs. The position concludes that since neither outcome is morally deserved, there is a *prima facie* reason to redistribute (some of) the fruits of the former to offset the negative consequences of the latter. Although scholars dispute the exact extent to which luck plays its role in individuals' lives, as well as the extent to which we can reliably determine this or remediate the negative consequences of bad luck, it is hard to dispute that luck plays at least some role in everyone's life.¹⁷ The hard questions are what we should, and what we can, do about it.

Both Hume and Smith discuss the effects of luck, or what they call "fortune," on human life. Hume speaks, for example of the "instability of fortune" (1742a: 150) and claims that "human life is more governed by fortune than by reason" (1742b: 180). Smith's treatment is more extensive, as he dedicates more than a dozen pages in TMS to what he calls the "influence of fortune" on our actions (TMS: 94–108), and in particular its influence on how we judge both our own and others' actions.¹⁸ Neither Smith nor Hume have much to say regarding fortune's effects on the allocation of resources, or on personal wealth, but what they do say about fortune provides a kind of response to the luck-egalitarian worries about markets and commercial society.

Hume writes: "Prosperity is naturally, though not necessarily, attached to virtue and merit; and adversity, in like manner, to vice and folly" (1741b: 553). Smith makes a similar claim:

If we consider the general rules by which external prosperity and adversity are commonly distributed in this life, we shall find, that notwithstanding the disorder in which all things appear to be in this world, yet even here every virtue naturally meets with its proper reward, with the recompense which is most fit to encourage and promote it; and this too so surely, that it requires a very extraordinary concurrence of circumstances entirely to disappoint it. What is the reward most proper for encouraging industry, prudence, and circumspection? Success in every sort of business. And is it possible that in the whole of life these virtues should fail of attaining it? Wealth and external honours are their proper recompense, and the recompense which they can seldom fail of acquiring. (TMS: 166)

¹⁵ For a fuller discussion of how a commercial society might conduce to a just and humane society, see Otteson (2019).

¹⁶ For a review of various positions, see Anderson (1999); see also Cohen (2009).

¹⁷ See Otteson (2017).

¹⁸ For discussion of Smith's position on "fortune," see Hankins (2016).

Recall that the central elements of Hume's and Smith's argument for the "humanity" of commerce were that it encourages proper moral relations among people, that it softens our prejudices and enlivens our virtues, and that it leads to increasing individual and overall prosperity. They both acknowledge that luck, or fortune, plays a role in each individual's life, and both of them recognize that we often fail to judge luck properly: we tend to praise or blame people for the consequences of their actions, which are often dependent on or affected by factors outside their control, instead of praising or blaming people only for their intentions. This seems like a moral mistake. Smith calls it an "irregularity of sentiments" (TMS: 106). As Smith argues, however, this "irregularity" has its beneficial consequences: because we know that others will judge us, however improperly or unfairly, on the effects of our actions and not just our intentions, it inclines us not only to take action but also to be much more deliberate in our actions, to take much greater precaution when considering our decisions than we otherwise might. If one knew instead that people judged one only on one's intentions, one might be "satisfied with indolent benevolence," or "fancy himself the friend of mankind, because in his heart he wishes well to the prosperity of the world" (ibid.). For Smith, however, this is not enough. To provide the benefit both to ourselves and to others of which we are capable, one must "call forth the whole vigour of his soul, and strain every nerve, in order to produce those ends which it is the purpose of his being to advance" (ibid.).

This gives us a new aspect of the morality, or the "humanity," of Humean/Smithian commercial society. What we might call "honorable business," or the summoning of our time, talent, and treasure to create value for ourselves only by simultaneously creating value for others, is what is rewarded both by the market *and* by human sentiment. As we have seen, according to Hume and Smith, in a "well-governed society," the only way we can do this is by recognizing the "opt-out option" of others, thus by treating them as our moral equals, and thus by respecting their agency such that it includes respect of their decisions—even when they say "no, thank you" to us, even when we believe they are mistaken, even when they are in fact mistaken. The commercial society holds out the prospect, then, of reward for taking positive steps to generate value for others. Though one "may often fulfil all the rules of justice by sitting still and doing nothing" (TMS: 82), and thereby would incur no resentment or cause for punishment, nevertheless the life of innocent inaction is not enough. We are capable not only of not injuring others but also of positively benefitting others. Using, within the bounds of justice, all the resources we have to do so is what can both create prosperity and lead us to enjoy the good judgment of others—something that both Hume and Smith claim we naturally desire. It also, according to both Hume and Smith, is the path to virtue and will be rewarded, not only in the next life but also in this one. We may speculate about what exactly the reward might be in the next life (or indeed whether there is a next life), but for Hume and Smith, the reward of virtue and prosperity in *this* life is not to be disparaged—or, apparently, doubted.

This would mean that the commercial society envisioned by Hume and Smith is one that encourages growth in "industry, knowledge, and humanity," that leads to proper respect of all persons' interests regardless of station, that is rewarded by one's own virtue

and thus happiness, and that, finally, is likely to be rewarded in this world if not in the next.¹⁹ The mandate to the individual lucky enough to live in a “well-governed society,” then, is, first, not to commit injustice, meaning not to infringe on others’ persons, property, or promises; and then, second, to actively seek ways to use one’s scarce resources to benefit oneself only by benefitting others. In a properly configured commercial society, this means engaging in honorable business, which seeks only cooperative and mutually beneficial ways to engage with others and respects all parties as persons of dignity whose decisions demand respect. In this way, the honorable businessperson contributes to a society that is both just and humane. And in this way, according to Hume’s and Smith’s vision of humane political economy, honorable business can indeed be a moral calling.

REFERENCES

- Anderson, Elizabeth. 1999. “What Is the Point of Equality?” *Ethics* 109: 287–337.
- Barber, Benjamin. 2008. *Consumed: How Markets Corrupt Children, Infantilize Adults, and Swallow Citizens Whole*. New York: Norton.
- Brennan, Jason, and Peter M. Jaworski. 2016. *Markets without Limits: Moral Virtues and Commercial Interests*. New York: Routledge.
- Clark, Gregory. 2007. *A Farewell to Alms: A Brief Economic History of the World*. Princeton, NJ: Princeton University Press.
- Cohen, G. A. 2009. *Why Not Socialism?* Princeton, NJ: Princeton University Press.
- Conly, Sarah. 2013. *Against Autonomy: Justifying Coercive Paternalism*. New York: Cambridge University Press.
- Conly, Sarah. 2015. *One Child: Do We Have a Right to Have More?* New York: Oxford University Press.
- Fleurbaey, Marc. 1995. “Equal Opportunity or Equal Social Outcome?” *Economics and Philosophy* 11: 25–55.
- Frankfurt, Harry G. 1998. “Equality as a Moral Ideal.” In *The Importance of What We Care About*. New York: Cambridge University Press.
- Gintis, Herbert. 2012. “Giving Economists Their Due.” *Boston Review*. Available at <http://bostonreview.net/gintis-giving-economists-their-due>.
- Hall, Joshua C., and Robert A. Lawson. 2014. “Economic Freedom of the World: An Accounting of the Literature.” *Contemporary Economic Policy* 32, 1: 1–19.
- Hankins, Keith. 2016. “Adam Smith’s Intriguing Solution to the Problem of Moral Luck.” *Ethics* 126: 711–46.
- Hoffman, Mitchell, and John Morgan. 2013. “Who’s Naughty? Who’s Nice? Experiments on whether Pro-Social Workers Are Selected out of Cutthroat Business Environments.” Rotman School of Management Working Paper. Available at https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2345102.
- Hume, David. 1740. *A Treatise of Human Nature*. David Fate Norton and Mary J. Norton (eds.). New York: Oxford University Press (2000 edition).
- Hume, David. 1741a. “Of the Independency of Parliament.” In Eugene F. Miller (ed.), *Essays Moral Political and Literary*. Indianapolis, IN: Liberty Fund (1985 edition), pp. 42–6.

¹⁹ Hume writes: “the happiest disposition of mind is the *virtuous*” (1742b: 168; italics in the original).

- Hume, David. 1741b. "Of Impudence and Modesty." In Eugene F. Miller (ed.), *Essays Moral Political and Literary*. Indianapolis, IN: Liberty Fund (1985 edition), pp. 552–6.
- Hume, David. 1742a. "The Stoic." In Eugene F. Miller (ed.), *Essays Moral Political and Literary*. Indianapolis, IN: Liberty Fund (1985 edition), pp. 146–54.
- Hume, David. 1742b. "The Sceptic." In Eugene F. Miller (ed.), *Essays Moral Political and Literary*. Indianapolis, IN: Liberty Fund (1985 edition), pp. 159–80.
- Hume, David. 1754a. "Of Refinement in the Arts." In Eugene F. Miller (ed.), *Essays Moral Political and Literary*. Indianapolis, IN: Liberty Fund (1985 edition), pp. 268–80.
- Hume, David. 1754b. "Of the Jealousy of Trade." In Eugene F. Miller (ed.), *Essays Moral Political and Literary*. Indianapolis, IN: Liberty Fund (1985 edition), pp. 327–31.
- Landes, David. 1999. *The Wealth and Poverty of Nations: Why Some Are So Rich and Some So Poor*. New York: Norton.
- Mandeville, Bernard. 1723. *The Fable of the Bees: Or, Private Vices, Publick Benefits*. Indianapolis, IN: Liberty Fund (1988 edition).
- Marx, Karl. 1844a. "Comments on James Mill, *Éléments D'économie Politique*." Available at <https://www.marxists.org/archive/marx/works/1844/james-mill/>.
- Marx, Karl. 1884b. "Alienated Labor." In Lawrence H. Simon (ed.), *Karl Marx: Selected Writings* (Indianapolis, IN: Hackett, 1994 edition).
- McCloskey, Deirdre N. 2006. *The Bourgeois Virtues: Ethics for an Age of Commerce*. Chicago, IL: University of Chicago Press.
- McCloskey, Deirdre N. 2010. *Bourgeois Dignity: Why Economics Can't Explain the Modern World*. Chicago, IL: University of Chicago Press.
- McCloskey, Deirdre N. 2016. *Bourgeois Equality: How Ideas, Not Capital or Institutions, Enriched the World*. Chicago, IL: University of Chicago Press.
- Nussbaum, Martha. 2011. *Creating Capabilities: The Human Development Approach*. Cambridge, MA: Harvard University Press.
- Otteson, James R. 2013. "Adam Smith." In Roger Crisp (ed.), *The Oxford Handbook of the History of Ethics* (New York: Oxford University Press), pp. 421–42.
- Otteson, James R. 2017. "The Misuse of Egalitarianism in Society." *The Independent Review* 22 (1): 37–47.
- Otteson, James R. 2019. *Honorable Business: A New Framework for Business in a Just and Humane Society*. New York: Oxford University Press.
- Piketty, Thomas. 2014. *Capital in the Twenty-First Century*. Cambridge, MA: Belknap Press.
- Radin, Margaret Jane. 2001. *Contested Commodities: The Trouble with Trade in Sex, Children, Body Parts, and Other Things*. Cambridge, MA: Harvard University Press.
- Robbins, Lionel. 1932. *An Essay on the Nature and Significance of Economic Science*. London: Macmillan.
- Rousseau, Jean-Jacques. 1755. *Discourse on the Origin of Inequality*. In Victor Gourevitch (ed. and trans.), *The Discourses and Other Early Political Writings* (New York: Cambridge University Press, 1997 edition), pp. 111–222.
- Rousseau, Jean-Jacques. 1762. *Of the Social Contract*. In Victor Gourevitch (ed. and trans.), *The Social Contract and Other Later Political Writings* (New York: Cambridge University Press, 1997 edition), pp. 39–153.
- Sandel, Michael J. 2012. *What Money Can't Buy: The Moral Limits of Markets*. New York: Farrar, Straus and Giroux.
- Satz, Debra. 2010. *Why Some Things Should Not Be for Sale: The Moral Limits of Markets*. New York: Oxford University Press.

- Scheidel, Walter. 2017. *The Great Leveler: Violence and the History of Inequality from the Stone Age to the Twenty-First Century*. Princeton, NJ: Princeton University Press.
- Schor, Juliet B. 2004. *Born to Buy: The Commercialized Child and the New Consumer Culture*. New York: Scribner's.
- Sen, Amartya. 1992. *Inequality Reexamined*. Cambridge, MA: Harvard University Press.
- Skidelsky, Robert, and Edward Skidelsky. 2012. *How Much Is Enough? Money and the Good Life*. New York: Other Press.
- Smith, Adam. 1759. *The Theory of Moral Sentiments*. D. D. Raphael and A. L. Macfie (eds.). Indianapolis, IN: Liberty Fund (1982 edition).
- Smith, Adam. 1776. *An Inquiry into the Nature and Causes of the Wealth of Nations*. R. H. Campbell and A. S. Skinner (eds.). Indianapolis: Liberty Fund (1981 edition).
- Thaler, Richard H. 2015. *Misbehaving: The Making of Behavioral Economics*. New York: Norton.
- Thaler, Richard H., and Cass R. Sunstein. 2009. *Nudge: Improving Decisions about Health, Wealth, and Happiness*. New York: Penguin.
- Tomasi, John. 2012. *Free Market Fairness*. Princeton: Princeton University Press.
- Ubel, Peter A. 2009. *Free Market Madness: Why Human Nature Is at Odds with Economics—And Why It Matters*. Cambridge, MA: Harvard Business Review Press.
- Zak, Paul J. (ed.). 2008. *Moral Markets: The Critical Role of Values in the Economy*. Princeton, NJ: Princeton University Press.
- Zak, Paul J. (ed.). 2012. *The Moral Molecule: The Source of Love and Prosperity*. New York: Dutton.

CHAPTER 15

CAPITALISM AND DEMOCRACY

Allies, Rivals, or Strangers?

JULIAN REISS

15.1 INTRODUCTION

AFTER the collapse of the Soviet Union in 1991, for a time the debate between the relative virtues and vices of “socialism” and “capitalism” all but disappeared. It was actually existing socialism that had collapsed, not just some individual regimes for their own specific reasons. Capitalism was widely regarded as the superior system, and not just because it outperformed rival systems economically. It was also widely seen as politically superior, because, or so it was believed at the time, democracy could thrive only in a market economy (see, for instance, the contributions to Diamond and Plattner 1993).

A quarter century down the line, things do not seem all that rosy anymore. Owing to increased macroeconomic volatility, enormous inequalities in income and wealth (both within individual countries and between developed and least developed countries), a financially stagnating and numerically shrinking middle class, unsustainable levels of sovereign debt, and the resurgence of nationalism in Europe and the United States, the race has once more opened up.¹ No longer does capitalism, at least in its current form, seem to be superior, economically or morally, to alternative economic systems. According to some commentators, indeed, capitalism’s end is nigh and the question is only *how*, not *whether*, it will end (Streeck 2017).

¹ These claims could be supplemented by at least as large a number of claims that things are getting better: extreme poverty has fallen from 37 percent to 9.6 percent since 1990, life expectancy is way up, women have fewer and more healthy children, enormous material progress has been made both globally and within most nations, and the progress has often been most significant for the poorest. However, many researchers in political economy worry about the phenomena described above, which are the worries that motivate renewed interest in the foundations of capitalism.

It is high time then to revisit the foundations of our social and economic order. This chapter will focus specifically on the interrelations between democratic institutions on the one hand and alternative economic arrangements on the other. It will address the following questions:

- Does capitalism promote or prevent democracy (or vice versa), or are these independent phenomena?
- If the two are related, what explains their relationship? That is, what are the mechanisms by which capitalism promotes or prevents democracy (or vice versa)?
- What kinds of freedom play a role in this relationship (if there is any)?

Although the focus of this chapter will be on the interrelations between democratic and economic institutions, and not on conceptual explorations into the meaning of the terms “capitalism” and “democracy,” I do want to state very briefly what I mean by the terms.² I want to characterize both by a short list of features that may or may not be present in an individual instance of the concept rather than strict necessary and sufficient conditions. Thus, capitalism refers to an institutional arrangement that guarantees:

- a. Private, individual property rights (individuals have the right to acquire, use, transform, and sell or dispose of property in land and tangible goods);
- b. Individual economic freedom (individuals have “property in themselves” and therefore own their labor initially and have the freedom to move and to make contracts to buy and sell land, goods, and labor);
- c. The existence of money and debt (contractual obligations can be settled by money and property can be used as collateral in order to advance or defer payments in debt contracts).

Democracy, in turn, refers to an institutional arrangement in which:

- a. Decisions concerning certain matters are made collectively and in a way that is binding for the group;
- b. The collective decision-making process uses some version of the majority vote;
- c. Basic political and civil rights (which include, for instance, the right to be elected for office) are guaranteed.

Put this way, a tension is immediately visible. Capitalism is all about the individual, individual decision making, and individual rights; democracy, about the collective, collective decision making, and political and civil rights. Socialism can be said to resolve the tension by extending the principle of collective decision-making to decisions about productive resources such as land, labor, and other productive goods. This also means

² I prefer the term “the liberal economy” to “capitalism.” The latter is a term of abuse, invented, and still used for the most part, by its opponents. “Liberal economy” carries its virtues—the centrality of freedoms and the rights of the individual—on its sleeve.

that property rights now lie with the collective (or a government body representing the collective) and that money and debt are severely constrained, if not nonexistent.

In what follows, I will first discuss the answers to the questions raised above that were given by important classical liberal, Marxist, and conservative thinkers, and then survey some of the more recent work on the topic. Let us begin where any serious discussion of capitalism and democracy should begin—with Alexis de Tocqueville.

15.2 TOCQUEVILLE

I start with the most profound and difficult thinker of the bunch. Alexis de Tocqueville (1805–59) was a French lawyer and diplomat and, according to Jon Elster (2009), the world's first social scientist. By this, Elster means that unlike all social thinkers who preceded him (and many who came after him), Tocqueville did not develop a social utopia as normative ideal, but rather sought to explain the workings of existing institutions on the basis of observations, testimony by contemporary witnesses and other data. His magnum opus, *Democracy in America* (1835–40)—henceforth *DA*—is still unrivalled in the acuity and depth of its analysis and rightly regarded as a classic in political theory.

Tocqueville wrote at a transitional time when aristocracy had been abolished in the French Revolution but democracy had not quite been established yet in his native France. This experience, and having been born into a family of ancient nobility many of whose members have personally suffered in the revolutionary turmoils, provided him with a unique outlook on American institutions. America, after all, had never seen an aristocratic regime and was characterized by a fundamental equality between the members of its society right from the time of first settlements.

That Tocqueville's book teaches us about democracy is plain from its title. Less plain is that it contains excellent analyses of capitalism as well. The term "capitalism" was first introduced by the French socialists Louis Blanc and Pierre-Joseph Proudhon in the 1850s and 1860s and was thus unavailable to Tocqueville. However, he was as interested in America's *economic* arrangements as in its political arrangements and, indeed, he is sometimes portrayed as a remarkable economist (Swedberg 2009).

From Tocqueville's point of view, this is a natural conclusion. What he meant by democracy was essentially "equality of conditions" or "equality of status." In an aristocracy, an individual is born into a family of a certain social rank, and leaving their social rank is impossible for most. A democracy abandons this automatic assignment to social rank. However, individuals do seek to differentiate themselves from others and they can do so by wealth. It is therefore that they are interested in gaining wealth, industrious and actively in pursuit of betterment.³ Politics and economics are thus intertwined.

³ As an aside, Tocqueville, like Schumpeter but unlike classical and neoclassical economists since Adam Smith, believed that "greed" or "the desire for wealth" or "economic rationality" (important differences between these notwithstanding) are symptoms of the underlying social and economic arrangements

As Tocqueville famously observed, in a democracy people desire both equality *and* freedom; however, their passion for equality is stronger and more lasting than that for freedom (*DA*: Vol II, Part 2, ch. 1). As a liberal, he was more concerned about the protection of freedom, and so a fundamental question raised by him is the following: How can freedom be preserved in the light of democratic equality?

Democracy threatens freedom through three routes: materialism, individualism, and centralization. Democracy encourages materialism because wealth is a means for social differentiation, as just mentioned. In an aristocratic society, those who are well off do not have to fear losing their wealth, but so material well-being is not the goal of life. But nor do the poor think of material well-being because they cannot hope to make any significant advances (*DA*: Vol II, Part 2, ch. 10). It is equality of conditions that creates a taste for material well-being.

There is thus nothing wrong with a taste for material well-being as such, but it can be excessive. It often leads to a preference for orderliness, but excessive orderliness may end up in tyranny. The quest for material well-being can also lead to moral deterioration. And it can lead away from civic engagement. After all, if the goal is gain, the opportunity costs for voluntary work can be very high indeed.

To understand better why the latter two threaten freedom, it is useful to distinguish two aspects of freedom Tocqueville emphasized. There are, on the one hand, the institutionally implemented political and civic rights such as free speech, freedom of assembly, freedom of the press, and the right to vote, that Tocqueville calls “free institutions.” There is, on the other hand, what Tocqueville called the “spirit of liberty.” This includes the appreciation of certain values and the implementation of these values in concrete behaviors that turn legal principles into lived reality. A people cannot be called free, for example, if its members are constitutionally guaranteed the right to vote but few people actually vote or segments of the population are prevented by others to exercise their right. Similarly, a right to free assembly is worth nothing if owners of public houses will not rent out space to members of a party because they are threatened with violence should they do so.⁴ By undermining public engagement, striving for material advancement can thus threaten liberty because it undermines the spirit of liberty.

Individualism is perhaps the threat to freedom that is most closely linked to what we would call capitalism. By individualism, Tocqueville means an individual's withdrawal into his own family, abandoning society at large, rather than “egoism” or excessive self-love (*DA*: Vol II, Part 2, ch. 2). Individualism in this sense is only possible under capitalism because all alternative socio-economic orders make individuals more directly interdependent, most often through the social hierarchy. In this way, an individual is

rather than something that is ingrained into the human psyche independently of these arrangements. In contrast, Smith, for instance, maintained that humans have an inborn “disposition to truck and barter” (Smith 1776/2008: Book I, ch. 2).

⁴ James Schleifer (2012: 67–8) distinguishes the “passive” (institutional) from the “active” (behavioral) dimension of freedom. It is important not to confuse these with Isaiah Berlin's (1958) “negative” and “positive” freedom, however, which are *alternative* conceptions of freedom, whereas Tocqueville's are two mutually dependent and reinforcing aspects of the same thing.

tightly connected to the rest of society. Under capitalism, complete autarky is possible and all social collaboration voluntary. A corollary is that social bonds are weakened and individuals may lose sight of greater social purposes.

Individualism is a threat to freedom because it makes the individual vulnerable to exploitation by powerful bodies, most notably the state. A feudal lord protects his serfs not only from external enemies but also from potential abuses by higher nobility. This layer of protection is lost with democratic equality.

The final threat is also the most famous: centralization. Centralization puts people's liberty at risk by the accumulation of power. Where does power accumulate? In a democracy that tends to be in those institutions that are thought to represent the people: in the first instance the legislature, but also the executive and the administration. These branches of government could accumulate power, for instance, by assuming responsibilities that could be located at a lower level (such as in the township). The most famous threat Tocqueville saw was in the *tyranny of the majority*. This is a corollary of the power that lies with the legislature. If that is not limited, coalitions can quickly form that oppress the minority.

The main mechanisms by which these threats could be contained are decentralization and federalism, shifting political majorities, free associations (such as unions, companies, and business and civil associations) as counterbalances, religion, and constitutionally guaranteed rights of the individual. Democracy, in Tocqueville's view, could only work if it was significantly limited, and capitalist institutions such as property rights and firms could help to quench expansive democracy.

15.3 MARX

As Karl Marx (1818–83) is the best known of the thinkers I am surveying here, I will be relatively brief and only focus on the parts of his thought that are important to understand his views on the relationship between democracy and alternative economic systems. Specifically, I want to look at the basic tenets of Marxian economics (including his prediction of its demise), his theories of alienation and exploitation, and his notion of a “true” democracy.

Marx published the first volume of *Capital* (which contains the most important ideas of Marx's political economy) in 1867, just four years before Carl Menger's *Principles of Economics* (1871) and William Stanley Jevons' *Theory of Political Economy* (1871) came out. Both books, together with Léon Walras' *Éléments d'économie politique pure* (1874), initiated the “marginal revolution” in economics which would fundamentally change the way economists explain social phenomena and affect economic theorizing up to today. Timing could not have been worse, for Marx placed a theory of value in the center of his whole thesis about how the capitalist society works and the bourgeoisie exploits the working class that was not only already largely discredited at the time he wrote, but that was also just about to be replaced by a superior alternative.

According to the Labor Theory of Value, which Marx borrowed from the classical economist David Ricardo (1817), the value of a commodity is the amount of labor embodied in it, both directly through the labor that was needed (“socially necessary”) in its production and indirectly through the labor that was needed in the production of the capital goods that were used. The theory works, at best, under highly restrictive assumptions such as perfect competition and homogeneity of skills, and can be derived from the subjective theory of value the marginalists offered under these assumptions as a special case (see Schumpeter 1942: ch. 3).

Marx’s theory of exploitation, and thus the core of his critique of capitalism, rests on it, however. An individual is exploited if and only if he or she works more hours than are necessary to produce the goods he or she consumes. Under feudalism, it is easy to see how Marxian exploitation occurs: serfs had the right to till a plot on the open field of their lord’s manor, but in return had to work on the lord’s demesne for some days. It is not obvious that something analogous happens in capitalism, and Marx has to invoke the Labor Theory of Value to argue that it does.

Workers are forced to sell their labor power on the market because they lack access to the means of production. As there is competition, the wage rate is equal to the amount of labor necessary to produce it, which means the amount of labor necessary to produce sufficient goods to maintain a worker. This is the *exchange value* of labor. The capitalist, however, can extract from it its *use value*, its contribution towards the value of the goods for sale on the market. Use value exceeds exchange value because the worker produces more goods than are necessary to maintain him. The difference is called *surplus value*. Exploitation obtains because labor is the only source of value, and yet surplus value flows from worker to capitalist in the wage labor process. (The means of production such as machines and raw materials also contribute value but only at the rate of the costs to acquire or replace them.)

The second building block of Marx’s normative critique of capitalism is his theory of alienation. There are various facets of this theory (see, for instance, Elster 1985: 100–7), but one core idea is that the degree of the division of labor increases as capitalism develops. As a result, workers are forced to specialize more highly in a way that is likely to make their work repetitive and devoid of meaning. Where an artisan performs most steps of the production of a good himself, the modern factory worker is downgraded to becoming a mere appendage of the machinery. Human beings are fundamentally multifaceted individuals with many different innate capabilities. Modern factory work alienates workers in part because it prevents individuals from fully realizing their capabilities.

Marx, of course, also predicted that capitalism would come to an end. The main driving force behind its demise, according to him, was technological change. Competition incentivizes investment in labor-saving technology because it allows capitalists to produce more cheaply. However, since labor is the only source of value, accumulation of capital leads to a decrease in the rate of profit (which is surplus value divided by constant and variable capital). This is the famous “Law of Tendency for the Rate of Profit to Fall.” Marx’s central idea was that technological progress had a long-term

“labor saving bias,” which is in fact a very controversial assumption (see, for example, Roemer 1981). This process leads in turn to a concentration of business and, because of layoffs and the creation of an “industrial reserve army,” the increased immiseration of the working class (which is, to say the least, also highly controversial, as wage rates had started to *rise* rather than fall by the time Marx was writing *Capital*). The growth of misery and oppression will drive the masses to revolt and eventually bring about the replacement of capitalism by communism.

Marx was a critic and an advocate of democracy simultaneously, depending on what we mean by the term. He rejected liberal bourgeois democracy, which can only supply a delusive form of “abstract freedom” (given by political rights and the like). Because it is based on individualism, it contradicts, according to Marx, man as a social being and leads to atomization.

The society that would overcome the horrors of social atomization he called “true democracy.” True democracy would abolish the alienation between the individual and the political community by resolving the split between the egoistic interests of individuals in bourgeois society and the social character of political life. This is achieved by a higher unity where civil society and the state would cease to be distinct. This higher unity can, however, only be created when the economic grounds for the distinctions between civil society and state or between man and society can be overcome. These economic grounds are, of course, private property rights, and the resulting opposition between bourgeoisie and proletariat and egoism on both sides. True democracy can therefore only be achieved when (private) property rights are abolished and replaced by communal ownership. The notion of “true democracy” (which he used in earlier writings such as 1834’s *Critique of Hegel’s “Philosophy of Right”*) is thus not very different from what he would later call “communism” (Avineri 1968: 34).

15.4 SCHUMPETER

Joseph Alois Schumpeter (1883–1950) stands out among the commentators on capitalism and democracy in that he predicts the socio-economic order he favors to fail in the near future. As he explained in *Capitalism, Socialism, and Democracy* (1942), he thought that capitalism was essentially doomed. Not, as Marx had maintained, because of “inner contradictions in the logic of capitalism.” To the contrary, capitalism “works” and would, if left to its own devices, continue to increase material well-being for a long time. Instead, capitalism affects the mores of the people living under it—in Tocqueville’s terms, it affects the *spirit of liberty*. Capitalism turns on itself because it creates the conditions under which people become more critical of capitalism and eventually will call for a socialist order to be implemented. It is therefore that Schumpeter gives a great deal of thought to the question whether socialism is compatible with democracy.

To understand why Schumpeter predicted the downfall of capitalism, it is useful to examine briefly what he thought was essential to it. Schumpeter was, it is important to

recall, as much a critic of contemporary mainstream economics as he was of Marxian economics. In particular, he complained that neoclassical economists focused on perfect competition, thus pretending that this was the rule rather than the exception, and on a state of static equilibrium instead of the process of change (ibid.: 77ff). He maintained that the latter was essential to capitalism. Specifically, to him it was a process of *creative destruction*, in which entrepreneurs have great ideas for new goods, production processes, forms of transportation and communication, distribution channels, and so on, and get these ideas funded through banks and marketed. Innovations compete with existing products and processes, and, when superior, destroy the latter. A corollary of this “entrepreneur-driven innovation” view of capitalism is a much more optimistic take on monopoly and monopolistic practices such as patenting, industry secrecy and long-term contracts because they are necessary to protect innovations (ibid.: ch. 7).

Schumpeter thus did not see any economic reasons for this process to stop—any obstacle could be circumvented by further innovation. Like Tocqueville, Schumpeter believed that fundamental socio-economic arrangements have an effect on how people think and what people value. Capitalism encourages rationality and critical discourse.⁵ That is a good thing at first and helps to bring down old aristocratic structures and religious prejudice, but eventually the critical spirit will not spare capitalism itself and contribute to its demise.

There are a number of contributing factors. In developed economies, the innovative function is often not performed by entrepreneur-business men, but instead by employees in the research-and-development departments of large companies. This expropriates the bourgeoisie in a sense as former entrepreneurs now receive labor income, undermining their role and the social position of the bourgeoisie in society. Capitalist evolution would also destroy what Schumpeter called the “protecting strata” of society, the remnants of the old aristocracy that supplied qualified individuals to serve in offices of state, officer the army, and devise policy. However, “without protection by some non-bourgeois group, the bourgeoisie is politically helpless and unable not only to lead its nation but even to take care of its particular class interest” (ibid.: 138).⁶ A further factor is the undermining of its own institutions of small- and medium-sized business through competition, and because of that weakening of the very idea of private property and freedom of contract. Finally, capitalism creates a “new class” of intellectuals critical of capitalism (ibid.: ch. 13), and it leads to a disintegration of the bourgeois family (as leading a traditional family life becomes increasingly costly for its members; see ibid.: ch. 14).

⁵ I say “encourage” rather than “create” because Schumpeter admits that pre-capitalist economic activity furthers rationality to some extent. However, only with capitalism does monetary accounting play an important role in society and only with capitalism are large enough fortunes created to be attractive to the best brains. Both phenomena contribute to a rationalization of society.

⁶ On this point Schumpeter would agree with the Victorian social and cultural critic Matthew Arnold (1822–88), who also argued that the bourgeoisie was not fit for political office (Arnold 1861). Unlike Schumpeter, though, Arnold thought that this problem could be solved by providing a better education to the middle classes.

After addressing the economic viability of socialism, which he answers in the affirmative, Schumpeter asks about the relationship between socialism and democracy. He rejects the Marxian analysis, according to which, as we have seen, socialism is a necessary condition for democracy because only socialism overcomes the economic and ensuing political power of the capitalist class, and argues instead that economic arrangement and political organization are fundamentally independent. A main building block of his analysis is a characterization of democracy according to which “the democratic method is that institutional arrangement for arriving at political decisions in which individuals acquire the power to decide by means of a competitive struggle for the people’s vote” (ibid.: 269). For democracy to flourish, the “human material of politics” should be of high quality, the effective range of decisions should not be extended too far, government must command a well-trained bureaucracy with a strong sense of duty and *esprit de corps*, and there must be “democratic self-control” (for instance, there should be acceptance of government decisions even on the part of those who voted for the opposition candidate).

Historically, democracy was a product of capitalism and capitalism promotes the good functioning of democracy because, for instance, property rights provide a natural limit to the reach of government. However, for Schumpeter this did not mean that socialism cannot be democratic. His main idea is a separation between public *economic* and *political* management. Economic decisions (for instance, about the relative prices of goods) can be made by a technical bureaucracy without political interference. According to Schumpeter, this might mean, at least potentially, a *smaller* degree of politicization than he observed in his contemporary capitalist countries (many of which had nationalized or heavily regulated monopolistic industries).

With hindsight, Schumpeter’s predictions seem overly pessimistic with respect to capitalism and overly optimistic with respect to the economic viability of socialism as well as the compatibility of socialism and democracy. After all, over 75 years on, capitalism, though widely loathed and criticized especially by the intellectual class, still thrives and indeed has elevated over a billion people out of extreme poverty since 1950s, and socialism is still looking for a single viable and democratic instance. However, it is of course possible that capitalism just needs a little more time to undermine itself, and that there is no instance of an economically viable and democratic socialist country does not mean that this is not possible. Hayek has provided a number of arguments to the effect that there are deeper, structural reasons for these empirical correlations.

15.5 HAYEK

Friedrich August von Hayek (1899–1992) was an Austrian like Schumpeter and also studied at the University of Vienna with Friedrich von Wieser. But unlike his older fellow countryman, Hayek believed in neither the inevitability of the decline of capitalism, nor in the economic viability of socialism, nor in the compatibility of socialism and democracy.

To understand his stance on socialism, it is useful to recall a debate that raged in the 1920s and 1930s about the feasibility of economic calculation under socialism. The debate started with an article written by Ludwig von Mises (1920) in which he argued that because socialism means state ownership of the means of production, there can be no markets; with no markets there can be no (non-arbitrary) prices; and without (non-arbitrary) prices, there could be no rational allocation of the factors of production. Socialism thus abolishes the mechanism responsible for societal economizing behavior.

In the ensuing “Socialist Calculation Debate,” prominent socialist economists such as Oskar Lange, Abba Lerner, Fred M. Taylor, and Maurice Dobb attempted to refute von Mises’ logic and demonstrate the feasibility and indeed superiority of socialism over capitalism. The so-called “Lange model,” for instance, uses the mathematical resources of neoclassical economics to portray an economy in which a central planning board allocates investment and capital goods and labor and consumer goods are allocated by markets; the planning board *simulates* a market in capital goods by a trial-and-error process (Lange 1936).

Hayek made a decisive contribution to the debate in his 1945 essay “The Use of Knowledge in Society,” which argued that the economic problem is not one of finding a vector of relative prices given knowledge of everyone’s preferences and available means. The problem is rather how best to make use of the information (about preferences and available means) that is given not to any single mind but is instead dispersed across all members of society. The question, according to Hayek, is not whether there is any economic planning as all economic activity involves planning; it is whether the planning should be done by many individuals, all of whom have access to a small slice of the information that exists in society, or rather a single central planner. Because for most economic decisions the local knowledge of the particular circumstances in which individuals find themselves is crucial, a central planning system will never outperform a system which builds on decentralized markets in which individuals make the best use of their knowledge through the price system:

Fundamentally, in a system where the knowledge of the relevant facts is dispersed among many people, prices can act to coördinate the separate actions of different people in the same way as subjective values help the individual to coördinate the parts of his plan. (Ibid: 526)

The price system is, thus, a mechanism for communicating information that would not be available in the absence of decentralized markets. It is therefore that models such as Lange’s fail to get to the core of the issue: they assume information to be “given” which simply would not be available outside of a market system. Lange (and before him, Enrico Barone, in whose work Schumpeter grounds his belief in the economic viability of socialism) only shows that if all the relevant facts were known to a single planner, the social problem could be solved; he does not show how a solution is produced in a world in which each individual possesses only incomplete, partial information.

Hayek also differed from Schumpeter in his assessment of the compatibility of socialism and democracy. As we have seen, Schumpeter's view that the two are independent and compatible was based on a very thin conception of democracy. He rejected what he called the "classical doctrine" of democracy according to which democracy "realizes the common good by making the people itself decide issues through the election of individuals who are to assemble in order to carry out its will" (1942: 250). Schumpeter did not think there was such a thing as the common good or a general will.

Hayek agreed but provided an original argument and arrived at the opposite conclusion with respect to the compatibility issue. It is worth quoting him at some length:

The "social goal," or "common purpose," for which society is to be organized is usually vaguely described as the "common good," the "general welfare," or the "general interest." It does not need much reflection to see that these terms have no sufficiently definite meaning to determine a particular course of action. The welfare and the happiness of millions cannot be measured on a single scale of less and more. The welfare of a people, like the happiness of a man, depends on a great many things that can be provided in an infinite variety of combinations. It cannot be adequately expressed as a single end, but only as a hierarchy of ends, a comprehensive scale of values in which every need of every person is given its place. To direct all our activities according to a single plan presupposes that every one of our needs is given its rank in an order of values which must be complete enough to make it possible to decide among all the different courses which the planner has to choose. It presupposes, in short, the existence of a complete ethical code in which all the different human values are allotted their due place. (Hayek 1944: 398–9)

Although common ethical standards affect choices, they do not determine them. "Thou shalt not steal" may make me pay for the goods I chose but does not tell me what to choose in the first place. Even highly specific maxims such as "always buy fair trade" do not tell me whether to prefer coffee to tea or vice versa—and it is highly unlikely that they would command universal agreement.

Capitalism, because of its decentralized markets, makes do without what Hayek called a "complete ethical code." I may feel that "coffee is more important than tea" but I do not need anyone to agree with me. The central planner, however, must make a decision one way or another. And either way, the planner will come down on the coffee drinkers' or the tea drinkers' side.

Suppose, then, that democratic socialism is attempted. The problem is that because there is no complete ethical code, there will not be agreement on most issues. Majority vote is impractical unless the number of alternatives is small. But the number of alternatives is monstrous.⁷ Now, perhaps we can make use of Schumpeter's proposal and leave such technical issues as the relative prices for breakfast beverages to a bureaucracy. Even

⁷ To stick with the breakfast beverage example, suppose there are only four options. Even if the ranking was purely ordinal (that is, there is no intensity with which people prefer coffee to tea or vice versa) there would be $4 \times 3 \times 2 \times 1 = 24$ alternative rankings. But there are many more breakfast beverages and many things people value other than breakfast beverages.

if so—and this does not seem to be a feasible solution⁸—the potential for widespread disagreement remains if we try to decide over a more abstract and far-reaching goal such as “social justice.” Schumpeter did not problematize the choice of a social goal and began his defense of the viability of socialism with the assumption that everyone should get the same share of the national product. But surely there are many standards of justice and there is no reason to suppose that any of them would have a chance to achieve widespread agreement.

Hayek then argues that because individuals cannot agree on an economic plan—and neither can their democratically chosen representatives—more and more people come to believe that “if things are to get done, the responsible authorities must be freed from the fetters of democratic procedure” (1944: 423). Socialism leads to totalitarianism because individuals cannot agree on social value, and after a period of parliamentary squabble authoritarian leaders will appear who force the interests of some group on everyone else.

Moreover, even if these problems could be solved, perhaps because, counterfactually, there is widespread agreement on some social goal and everyone accepts the decisions of an economic planning bureaucracy, the resulting democracy would at best be one in Schumpeter’s sense (where there is genuine competition for leadership) because basic civil rights could not be guaranteed to everyone. Let us assume that the social goal is an equal wage for everyone. Since there are more and less attractive jobs, people would flock into the attractive ones, risking that many, or even most, things are left undone. But the things the planner determines to be socially valuable cannot be left undone and so the freedom of occupational choice must be restricted. Same-sex wedding cakes will either be available or not, thus either violating some bakers’ right to religious belief (because they will have to produce the cake whether they want to or not) or same-sex couples’ rights. More generally, whenever different groups’ preferences or interests clash, the government will have to come down on one side or the other, which will often imply a violation of the losing group’s rights.

Thus, as long as “democracy” implies a certain amount of civil liberties and the exercise of these liberties has economic implications—which it almost always does, because religious treatises must be printed, political speeches broadcast, associations meet somewhere, and so forth—democracy is difficult to combine with economic planning, which, in turn, is an essential ingredient in socialism. It is therefore that Hayek writes: “If ‘capitalism’ means here a competitive system based on free disposal over private property, it is far more important to realize that only within this system is democracy possible” (1944: 429).

While Hayek was certainly as worried as Schumpeter about capitalism’s reputation among the intellectuals and the general public, he did not predict its imminent end.

⁸ Schumpeter hoped that leaving technical issues to a bureaucracy would make economic decision-making less politicized. However, there is little reason to suppose it would. It seems unfair to tea drinkers if someone made a conscious decision to make coffee more easily available and vice versa. Eventually, the bureaucracy would come under immense political pressure.

Since in his view collectivism (which includes not just socialism but also fascism and Nazism) would necessarily lead to totalitarianism, it had to be fought and capitalism defended.

Finally, it is important to point out that Hayek, while sympathetic to democracy, did not see any intrinsic value in it. What was important to him was liberalism, the protection of freedoms, and the rule of law, which could be achieved in a democracy, but there is nothing desirable about democracy as such: “However strong the general case for democracy, it is not an ultimate or absolute value and must be judged by what it will achieve” (1960: 170). Democracy can be an adequate means to select among possible liberal laws, but it clearly had to be limited in order to be consistent with liberal principles.

15.6 POLANYI

Karl Polanyi (1886–1964) was Austrian by birth and, like Schumpeter and Hayek, spent an important part of his life in the United States. Between 1924 and 1933 he worked as an editor for the bi-weekly magazine *Der Oesterreichische Volkswirt*, but he was by no stretch of the imagination an economist of the Austrian School. In his editorial work he criticized the works of von Mises and Hayek as abstract and out of touch with the socially interrelated reality of economic processes.

Polanyi’s magnum opus *The Great Transformation* was published in 1944, the same year as Hayek’s *The Road to Serfdom*. While both are undoubtedly epoch-making books, they could not be more different in style and conclusion. *The Road to Serfdom* deals mainly in ideas and presents one of the most powerful defenses of free-market capitalism to date; *The Great Transformation* was steeped in historical analysis and ethnographic study and constitutes an equally powerful critique of market liberalism. It is well worth reading both side by side.

There are three terms that are central to Polanyi’s theory of market society: embedding, fictitious commodities, and the double movement. Polanyi observes that prior to the advent of capitalism, “the economic system [was] run on noneconomic motives” (1944: 48); economic activity was subordinate to other social practice. Polanyi uses the term *embedding* for this relationship between the economy and society. In pre-capitalist (tribal, feudal) orders, the economy is firmly embedded in society. This changed with the Industrial Revolution and, with it, the appearance of the “self-regulating market system” in the nineteenth century. According to Polanyi, the self-regulating market requires an institutional separation of society into an economic and a political sphere. This, in turn, functions only if society is subordinated to the market: “A market economy can exist only in a market society” (ibid.: 74).

Polanyi says this because in a market society, labor, land, and money are treated as commodities and therefore made subject to market forces. But in fact, labor is constituted by human activity which flows from life itself, land by the nature that surrounds

them, and money is a token of purchasing power which comes into being through banking or state action. They are *fictitious commodities* because they are not produced for sale in the market but treated as though they were. The problem with this is that if left to itself—and the market society can only function in this way—society will be destroyed. If market forces determine the value of labor, because labor is necessarily attached to human beings, the physical, psychological and moral integrity of individuals is at risk, which can lead to dislocation, disease, crime, starvation, and death. Treating land as commodity risks environmental degradation. Treating money as commodity risks periodical liquidation of businesses *en masse* because of currency or interest rate fluctuations. What is important, a self-regulating market that treats labor, land, and money in this way is not, pace Hayek (1960: 160), a “spontaneous order.” The market has been the outcome of a conscious intervention by government: “Laissez-faire was planned, planning was not” (Polanyi 1944: 147).

The *double movement* describes the dual tendency of increasing marketization on the one hand, and attempts to protect especially labor, land, and money from the worst consequences of marketization on the other. Because of the dangers of markets to destroy what is essential to man, nature, and money, a countermovement arose that served to check the action of the market through government interventions. Each arm of the double movement is supported by an ideology the main function in society of which is the bolstering of certain aims. Economic liberalism uses laissez-faire and free trade to strengthen marketization; what Polanyi calls “the principle of social protection” aims at the conservation of man, nature, and productive organization, and uses protective legislation, restrictive associations, and other instruments of intervention to achieve it (1944: 138–9).

Polanyi, like Schumpeter and Hayek, thought that modern democracy was a product of capitalism (*ibid.*: 74; 231). More than that, he thought that if capitalism disappears, so too will democracy. The movements of fascism and totalitarian socialism in the 1930s were reactions to the disintegration of the capitalist world order (*ibid.*: ch. 20). But this disintegration was not caused by any external force. The self-regulating market was stopped by the reality of society that does not accept the treatment of labor, land, and money as commodities, and was an inevitable result of the separation of the economy from society (or, alternatively, from the subordination of society to the economy). At this point there are choices. To accept the reality of society means to accept a modicum of coercion: “Power and compulsion are part of that reality; an ideal that would ban them from society must be invalid” (*ibid.*: 267). Power and compulsion can be glorified, as they have been in fascism and totalitarian socialism. But they can also be used for good. Through planning, control, and regulation, the rights of individuals can be strengthened and thus freedom for all achieved.

What Polanyi was advocating was essentially a form of welfare state in which markets would operate but in a regulated fashion and under democratic control; society would thus run the economy rather than vice versa, and the two would form an organic whole (as they used to prior to the Industrial Revolution). Roosevelt’s New Deal went a long way towards this ideal. Polanyi hoped that the threat of tyranny from the centralization

that comes along with the welfare state could be curbed by constitutionally guaranteeing spheres of arbitrary freedom to individuals.

Are capitalism and democracy compatible according to Polanyi? Ultimately, no. If by capitalism we mean the self-regulating market system, that order will collapse at the reality of society and call for something new. We can choose to preserve freedom and democracy in the new order, but the new order will differ from unfettered, nineteenth-century capitalism in important ways.⁹

15.7 THE POSTWAR LITERATURE

So far we have canvassed Marxist arguments to the effect that the capitalism and democracy are incompatible or “rivals.” Then there are two kinds of arguments that the two are “allies.” On the political right are arguments to the effect that capitalism is necessary for democracy (Hayek), and that democracy has to be limited for it to work (Tocqueville, Hayek). On the political left are arguments to the effect that capitalism and democracy can be mutually enforcing so long as democracy serves to limit capitalism (Polanyi). Finally, there are arguments to the effect that there are no necessary or causal relationships between the two; that they are “strangers” (Schumpeter).

There has been an explosion of literature in political economy since the Second World War that focuses mostly on the interaction of democratic institutions and the welfare state.¹⁰ Here I will focus on three bodies of literature that contain what I regard as the most exciting developments in this area: one neo-Polanyian view, one neo-Marxian view, and one neo-Tocquevillean view.

15.7.1 Varieties of Capitalism

After the collapse of the Soviet Union and its vassal states, there was a brief period of near consensus among political economists that (a) capitalism is economically superior to socialism; (b) that capitalism is necessary for democracy, at least in the sense that there has never been a case of a democracy that was not built on a market economy;¹¹ and (c) the welfare state supports the alliance between capitalism and

⁹ Polanyi observes, for instance, that “The nature of property, of course, undergoes a deep change in consequence of such measures since there is no longer any need to allow incomes from the title of property to grow without bounds, merely in order to ensure employment, production, and the use of resources in society” (1944: 260). Given that property rights are essential under any conception of capitalism, this change can be quite consequential.

¹⁰ For a survey, see Iversen (2006).

¹¹ This distinction is important: just because no case of democracy without capitalism has been observed, does not mean that it is impossible and will not emerge at some point in the future. Lindblom (2001), for

democracy by protecting those without property through redistribution and regulation (see, for instance, the contributions to Diamond and Plattner 1993). In this context, it is understandable that a number of political economists and sociologists turned away from the “big questions” concerning capitalism and socialism and instead examined differences in the institutional characteristics *among* different welfare capitalist nations. This is the “varieties of capitalism” (VoC) approach.

One contribution of the VoC literature is a typology. Rooted in Polanyi’s work, Gøsta Esping-Andersen (1990, 1999), for instance, distinguishes a liberal welfare state, which tends to minimize the role of the state in keeping with a market-oriented political culture (such as in the United States); a Nordic welfare state, which tends to provide comprehensive government coverage of risks, a generous level of benefits, and egalitarian tax policies (as in the Scandinavian countries); and a conservative welfare state, which shows the influence of Catholic social teaching (such as Germany and some Mediterranean countries). William Baumol et al. (2007) distinguish four different types of capitalism: entrepreneurial, big-firm, state-directed, and oligarchic.¹² Peter Hall and David Soskice (2001) distinguish a liberal market economy, a coordinated market economy, and a “Mediterranean” type, though in their analysis focus mainly on the former two (see also Hancké et al. 2007).

In my view, the most significant results in the emerging VoC literature have to do with the analysis of the interrelations among the economic and the political institutions of welfare states. Torben Iversen (2005: ch. 4), for instance, has shown that because proportional-representation electoral systems tend to promote left party dominance and redistribution, it enables workers to invest in specific skill acquisition in economies that rely heavily on workers with industry-specific skills (that is, coordinated market economies). The opposite is true in liberal market economies. Here, majoritarian electoral systems, leadership-dominated parties, small welfare states, and the development of general skills cluster.

The wider significance of the VoC approach for the “capitalism and democracy” issue is that it provides far more detailed and specific arguments for how different socio-economic and political institutions interlink in order to create a form of capitalism that is sustainable and successful (see, for instance, the explicit reference to Polanyi in Estevez-Abe et al. 2001).¹³

instance, argues that the reason for why no existing democracy has been transformed peacefully into democratic socialism is that, in a capitalist democracy, pro-market forces use their power to talk the masses into continuing to accept the market system, even if it is no longer in their interest to do so. Hayek, of course, has given us reasons to believe that as goes the market, so goes democracy.

¹² Their focus is on growth, however, and not on the interrelations between economic and political institutions.

¹³ There is a potential snag: Wolfgang Merkel (2014) argues that capitalism has changed dramatically after 1980, with its turn towards neoliberalism, deregulation and globalization, and the rise of financialization, in a way that undermines the compatibility of (financial) capitalism with democracy. Not surprisingly, he argues that capitalism must be re-embedded in order to avoid the teardown of democracy.

15.7.2 The Crises of Democratic Capitalism

Wolfgang Streeck (2017) has no patience for the kind of institutional subtlety the VoC literature offers. He believes that it masks important elements of capitalism that are common among the different varieties such as power differentials between owners and non-owners of capital and the resulting divergence of interests. Streeck agrees with the proponents of the VoC approach that the first three decades or so after the Second World War were characterized by shared prosperity and some democratic control of the economy through regulation and redistribution. The era was marked by a “peace formula” according to which the working classes accepted capitalist markets and property rights in exchange for political democracy, which in turn enabled them to achieve social security and a steadily rising standard of living (ibid.: 78). He sees “competition of systems” during the Cold War period as a distinct advantage for capitalism because “Socialism and trade unionism, by putting a brake on commodification, prevented capitalism from destroying its non-capitalist foundations—trust, good faith, altruism, solidarity within families and communities, and the like” (ibid.: 60). The three decades after the war were, however, the exception rather than the rule; the following series of crises represent the normal condition of democratic capitalism.

Unlike the Polanyians of the previous subsection, Streeck thus regards the financialization of capitalism not as a mere temporary move (back) towards *laissez-faire* liberalism, but instead as a manifestation of underlying contradictions that attempts at “embedding” will not overcome. The main tension embodied in democratic capitalism is the simultaneous adherence to two fundamentally different principles of resource allocation: one operating according to marginal productivity (or whatever the market rewards), and the other by social need or entitlement as determined by collective decision making. The crises of high inflation in the late 1970s, of unemployment in the early 1980s, of public debt in the late 1980s and 1990s, that of austerity and deregulation in the late 1990s and 2000s, and the financial crisis of 2008 were all an expression of this tension:

Toleration of inflation, acceptance of public debt and deregulation of private credit were no more than temporary stopgaps for governments confronted with an apparently irrepressible conflict between the two contradictory principles of allocation under democratic capitalism: social rights on the one hand and marginal productivity, as evaluated by the market, on the other. (Streeck 2017: 90)

As the fundamental tension is inherent in democratic capitalism, there are few reasons to believe that this sequence of crises in ever new variants is going to stop. Like Schumpeter (and Marx), Streeck predicts the (soon-ish) demise of capitalism after a number of further rounds of crises. However, he regards it as a Marxist prejudice that it should come to an end only when a new, better order is on the horizon (ibid.: 57). Capitalism will disappear when it no longer keeps its promise as a self-reproducing, sustainable, predictable, and legitimate social order, which, he argues, is well on its way. The reason for its demise is also Marxian: its inner contradictions.

15.7.3 Rent-Seeking

An economic *rent*, in the parlance of modern public choice theorists, is an excess return to a production factor beyond what would be needed to keep it in its current use. If, for example, a government decides to require a license for someone to drive a taxi and thereby reduces the number of taxi operators, it creates a rent because taxi drivers can charge higher prices than they could in a genuinely competitive environment.

Up until the late 1960s, mainstream economists believed that licenses and numerous other government interventions give rise to two problems. On the one hand, there is a transfer of wealth from consumers to the beneficiaries of the intervention such as our taxi drivers. On the other, there is what economists call a *deadweight loss*, the social cost of the intervention. The latter is created by the fact that there are numerous consumers who would pay for the service at a price between the higher, licensed price and the competitive price but cannot do so because licensing prevents it.

In a 1967 article, Gordon Tullock argued that the social cost of the intervention is even higher because the promise of rents induces producers to invest in activities (such as lobbying) intended to create rents in the first place. Anne Krueger (1974) coined the term *rent-seeking* for this kind of behavior. There are, thus, three adverse effects of an intervention such as licensing: wealth transfer from consumer to producer, a deadweight or social loss, and the incentivization of rent-seeking behavior (which is not regarded as productive).

This basic idea has been applied in a vast number of contexts, including: monopoly and regulation of industry; protectionist international trade policies; economic development; property rights and corruption (when the rule of law is weak); migration; electoral politics; the courts, the judiciary, and litigation; institutions (the rules of the game itself); alternative economic systems such as mercantilism and authoritarian regimes; and “soft budgets” created through government transfers. (For a collection of the most important work, see Congleton et al. 2008.) The upshot of this literature is that much well-intended government intervention and regulation can have severe adverse unintended consequences.¹⁴ Creating a monopoly carries social costs (in the form of deadweight or efficiency losses and rent-seeking) and distributive costs. The distributive costs can depend on the number of competitors for the rent. If that is large, the winner may incur costs that are as high as the rent, thus attenuating the distributive effect (because the expenses for rent-seeking are dissipated) and increasing the social costs. If the number of competitors is small, there can be considerable redistribution from bottom to top.

¹⁴ Cf. Frédéric Bastiat’s essay “What Is Seen and What Is Not Seen” (in Bastiat 1848/1995) where he uses the parable of the broken window to illustrate why destruction, and the money spent to recover from destruction, is not actually a net benefit to society. The money spent on fixing the broken glass (which is seen) does provide income to the glazer, but the shopkeeper whose glass was broken cannot spend the money on an alternative purpose (which is not seen). Similarly, regulation has effects that can be seen—such as that certain practices are no longer performed, certain groups’ incomes are increased, or “inferior” foreign goods are kept out of the market—but also a variety of unseen efficiency and distributive effects, many of which can be detrimental to the well-being of society.

There is also another effect that has to do with collective decision making, identified by Mancur Olson (1982). Members of small groups have disproportionate organizational power for collective action because they can co-ordinate their actions more easily than large groups. But this means that socially disadvantaged groups, which are often large and not well organized, are less able to influence economic policy in their favor:

The recipients of welfare in the United States are not organized, nor are the poor in other societies. But in the United States, as elsewhere, almost all the major firms are represented by trade associations and the professions by professional associations.

(Ibid.: 363)

Existing attempts to regulate and redistribute may therefore create greater inequality rather than less.

15.8 CONCLUSION

Let us return to Polanyi as we end this chapter. Polanyi argued that the commodification of labor, land, and money creates the greatest obstacle for unfettered capitalism to work. Only regulation that protects labor (through workplace safety, unemployment and social insurance, and redistribution), land (through environmental regulation, building codes, and so forth), and money (through repealing the gold standard, limiting capital movements and managing the currency) can achieve a re-embedding of the economy in (democratic) society and sustainable economic organization.

However, to the extent that public-choice theorists have got it right, the actual effects of regulation and redistribution will often be the opposite of the intended effects. The crises of capitalism Wolfgang Streeck and others describe may well be caused by the aggregated rent-seeking activities of numerous interest groups rather than capitalism as such. (For some considerations to that effect, see Reich 2015.)

We seem to be stuck between a rock and a hard place. Laissez-faire capitalism does not work (says Polanyi), but neither does embedded capitalism (say the public-choice theorists).¹⁵ Given the overwhelming empirical and theoretical reasons to believe that democracy cannot be had without capitalism, we can only hope that either Polanyi or the public-choice theorists are wrong or, perhaps, that there is a hitherto unnoticed form of capitalism that does work.

ACKNOWLEDGMENTS

I wish to thank Bruce Caldwell and Mark D. White for valuable comments on a previous version of this chapter. Funding from the European Research Council for the project “Knowledge for Use” (Grant agreement No. 667526 K4U) is gratefully acknowledged.

¹⁵ I ignore here orders that are based on a full-on nationalization of productive resources because they have been theoretically and empirically discredited.

REFERENCES

- Arnold, Matthew. 1861. "Democracy." In *The Popular Education of France* (London, Longman), pp. 1–47.
- Avineri, Shlomo. 1968. *The Social and Political Thought of Karl Marx*. Cambridge: Cambridge University Press.
- Bastiat, Frédéric. 1848. *Selected Essays on Political Economy*. Irvington-on-Hudson: Foundation for Economic Education (1995 edition).
- Baumol, William, Robert Litan, and Carl Schramm. 2007. *Good Capitalism, Bad Capitalism, and the Economics of Growth and Prosperity*. New Haven, CT: Yale University Press.
- Berlin, Isaiah. 1958. *Two Concepts of Liberty*. Oxford: Clarendon Press.
- Congleton, Roger D., Ayre L. Hillman, and Kai A. Konrad. 2008. *40 Years of Research on Rent Seeking 2: Applications: Rent Seeking in Practice*. New York: Springer.
- de Tocqueville, Alexis. 1835–1840. *Democracy in America*. New York: Library of America (2004 edition).
- Diamond, Larry and Marc Plattner (eds). 1993. *Capitalism, Socialism, and Democracy Revisited*. Baltimore, MD: Johns Hopkins University Press.
- Elster, Jon. 1985. *Making Sense of Marx*. Cambridge: Cambridge University Press.
- Elster, Jon. 2009. *Alexis De Tocqueville, the First Social Scientist*. Cambridge: Cambridge University Press.
- Esping-Andersen, Gøsta. 1990. *The Three Worlds of Welfare Capitalism*. Cambridge: Polity Press.
- Esping-Andersen, Gøsta. 1999. *Social Foundations of Postindustrial Economies*. Oxford: Oxford University Press.
- Estevez-Abe, Margarita, Torben Iversen, and David Soskice. 2001. "Social Protection and the Formation of Skills: A Reinterpretation of the Welfare State." In Peter A. Hall and David Soskice (eds), *Varieties of Capitalism: The Institutional Foundations of Comparative Advantage* (Oxford: Oxford University Press), pp. 145–83.
- Hall, Peter A., and David Soskice (2001). *Varieties of Capitalism: The Institutional Foundations of Comparative Advantage*. Oxford: Oxford University Press.
- Hancké, Bob, Martin Rhodes, and Mark Thatcher. 2007. *Beyond Varieties of Capitalism: Conflict, Contradictions, and Complementarities in the European Economy*. Oxford: Oxford University Press.
- Hayek, Friedrich. 1944. *The Road to Serfdom*. Chicago, IL: University of Chicago Press.
- Hayek, Friedrich. 1945. "The Use of Knowledge in Society." *American Economic Review* 35: 519–30.
- Hayek, Friedrich. 1960. *The Constitution of Liberty*. Chicago, IL: University of Chicago Press.
- Iversen, Torben. 2005. *Capitalism, Democracy, and Welfare*. Cambridge: Cambridge University Press.
- Iversen, Torben. 2006. "Democracy and Capitalism." In Barry R. Weingast and Donald A. Wittman (eds), *The Oxford Handbook of Political Economy* (Oxford: Oxford University Press), pp. 601–23.
- Levons, William Stanley. 1871. *The Theory of Political Economy*. New York: Palgrave (2013 edition).
- Krueger, Anne. 1974. "The Political Economy of the Rent-Seeking Society." *American Economic Review* 64: 291–303.
- Lange, Oskar. 1936. "On the Economic Theory of Socialism." *Review of Economic Studies* 4: 53–71.
- Lindblom, Charles. 2001. *The Market System: What It Is, How It Works, and What to Make of It*. New Haven, CT: Yale University Press.

- Marx, Karl. 1834. *Critique of Hegel's "Philosophy of Right."* Annette Jolin and Joseph O'Malley (trans). Cambridge: Cambridge University Press (1970 edition).
- Marx, Karl. 1867. *Capital: Volume One*. Ben Fowkes (trans.). New York: Penguin (1976 edition).
- Menger, Carl. 1871. *Principles of Economics*. Glencoe, IL: Free Press (1950 edition).
- Merkel, Wolfgang. 2014. "Is Capitalism Compatible with Democracy?" *Zeitschrift für vergleichende Politikwissenschaften* 8: 109–28.
- Olson, Mancur. 1982. *The Rise and Decline of Nations*. New Haven, CT: Yale University Press.
- Polanyi, Karl. 1944. *The Great Transformation: The Political and Economic Origins of Our Time*. Boston, MA: Beacon Press (2001 edition).
- Reich, Robert. 2015. *Saving Capitalism: For the Many, Not the Few*. New York: Alfred A. Knopf.
- Ricardo, David. 1817. *The Principles of Political Economy and Taxation*. Mineola, NY: Dover (2004 edition).
- Roemer, John. 1981. *Analytical Foundations of Marxian Economics*. Cambridge: Cambridge University Press.
- Schleifer, James. 2012. *The Chicago Companion to Tocqueville's Democracy in America*. Chicago, IL: University of Chicago Press.
- Schumpeter, Joseph. 1942. *Capitalism, Socialism, and Democracy*. New York: Routledge (2003 edition).
- Smith, Adam. 1776. *An Inquiry into the Nature and Causes of the Wealth of Nations: A Selected Edition*. Oxford: Oxford University Press (2008 edition).
- Streeck, Wolfgang. 2017. *How Will Capitalism End? Essays on a Failing System*. London: Verso.
- Swedberg, Richard. 2009. *Tocqueville's Political Economy*. Princeton, NJ: Princeton University Press.
- Tullock, Gordon. 1967. "The Welfare Costs of Tariffs, Monopolies, and Theft." *Economic Inquiry* 5: 224–32.
- von Mises, Ludwig. 1920. "Economic Calculation in the Socialist Commonwealth." In F. A. Hayek (ed.), *Collectivist Economic Planning: Critical Studies on the Possibilities of Socialism* (London: Routledge, 1935), pp. 87–130.
- Walras, Léon. 1874. *Elements of Theoretical Economics: Or, the Theory of Social Wealth*. Cambridge: Cambridge University Press (2014 edition).

CHAPTER 16

THE MORAL STATUS OF PROFIT

JOSEPH HEATH

MANY critics of capitalism consider it self-evident that the system is unethical because it allows, and indeed relies upon, the organization of economic activity based on the pursuit of profit. Unfortunately, because these critics regard the proposition as self-evident, they seldom take the time to specify in much detail why they think the pursuit of profit is so problematic. In some cases the animus seems to be based on an elementary confusion between the self-interest of individuals and the profit orientation of firms. If moral rules constitute impartial constraints on the pursuit of self-interest, and profit-maximization is just another way of describing the pursuit of self-interest, then it follows that morality must stand in some sort of antagonistic relation to the pursuit of profit.

This confusion, it should be noted, has not been confined to critics of capitalism. The undergraduate economics textbook that I studied in the 1980s started out with the assumption that individuals are self-interested, and then introduced utility functions in order to represent this (Lipsey et al. 1988). All of a sudden these individual utility functions were “aggregated” into joint utility functions to represent “households,” which were then taken to want to maximize consumption, and “firms,” which were taken to want to maximize profits (46). This made it seem as though “profit-maximization” on the part of the firm was just individual self-interest writ large. This represents a very unhelpful confusion. With the further development of economic theory, however, it became clear that this sort of aggregation is extremely problematic. Thus it has become standard in contemporary treatments to draw a clear distinction between the motivations of individual agents and profit as an *organizational objective*, which individuals may or may not have an interest in pursuing.

At the same time, there is a kernel of truth to the point about self-interest, which is that firms, in seeking to maximize profits, are exhibiting a form of *partiality*, rather than acting on the basis of an impartial concern for the interests of all. They are, for instance, not directly aiming to increase the sum of human happiness, but are instead privileging the welfare of various firm “insiders,” in particular, the ownership group. This sort of

partiality, however, is not uniformly forbidden by everyday morality—people act very much the same way toward their own family—and there are various conditions under which it can be perfectly compatible with an impartial concern for the interests of all. Thus it must be explained why this form of partiality is problematic in the specific case of the business corporation, and there have been a bewildering number of proposals (for example, see Giroux 2008: 2; Skidelsky and Skidelsky 2012). According to some, the pursuit of profit is what drives firms to externalize costs (Greenfield 2006), for others the earning of profits encourages the exploitation of labor (Boltanski and Chiapello 2005: 373), for others the competitive dynamic unleashed by the pursuit of profit erodes community attachments (Cohen 2009), and so on. Given this proliferation of claims, many articles that seek to defend the profit orientation of firms begin reconstructively, having to catalogue the various ways in which people have thought that profit is bad, before attempting to marshal a response (Flew 1976; Walsh and Lynch 2008: 66–74).

Rather than add my own two cents to this increasingly voluminous literature, I would instead like to focus on a peculiarity of the profit orientation that has often escaped notice. In particular, I would like to criticize a common fallacy: that because firms are licensed to pursue profits, the purpose or goal of the economic system as a whole must be to facilitate the realization of such profits. This is manifestly not the case, since under “ideal” market conditions firms would be unable to earn any profits at all. More concretely, the design of markets, and in particular the insistence on competition between firms, is intended to bid profits down to zero. This is because the pursuit of profit on the part of different firms operating in the same sector (that is, producing the same type of good) is *collectively self-defeating*. In seeking to increase their own profits, firms act in ways that undermine the profitability of the sector as a whole. This is a feature of the institutional design of markets. The point of allowing firms to pursue profits is precisely to induce this sort of collective action problem, as evidenced by the fact that cartels are illegal and, in sectors that are monopolized by one firm, government regulation typically prohibits profit-maximization. In the same way that two criminals, stuck in a prisoner’s dilemma, act dishonorably toward one another and yet wind up “doing the right thing” for society as a whole by turning each other in, profit-maximizing firms in a competitive market wind up acting in a way that precisely mirrors the prisoner’s dilemma. This means that, in terms of moral assessment, it is necessary to look at the net effect of their actions, when undertaken in the appropriate institutional context, in order to see whether they also are “doing the right thing” for society as a whole.

Economic theory provides a standard answer to this question. The point of allowing firms to maximize profits is that the lure of profit is what leads them to compete with one another. In a well-structured market, this competition drives prices toward the level at which supply is equal to demand (that is, at which “markets clear”), and these prices in turn allow for a more efficient allocation of labor, resources, goods and services across the economy as a whole. The latter is in fact the “point” of the system. Since rational economic decision-making is impossible without some system of prices, the alternative to market determination of prices is some form of explicit calculation, an exercise that has so far proven impossible in practice. This is why the great socialist economist Oskar

Lange once wrote that the central attraction of capitalism is that “competition forces entrepreneurs to act exactly as they would have to act were they managers of production in a socialist system” (1937: 123). From this perspective, the encouragement to pursue profit is essentially a trick, designed to induce firms to promote an end that is no part of their intention.

This analysis, which I will attempt to articulate at greater length in this chapter, explains why so many people find an economic system based on profit-maximization to be morally counterintuitive, but also provides the foundation for its ultimate vindication. It is counterintuitive because much of the force of everyday morality is aimed at getting people to act more cooperatively, whereas profit-maximization is essentially a free-rider strategy. The idea of a socially induced collective action problem is, at best, unfamiliar, and poses particular challenges from the standpoint of moral justification. The system of market competition is ultimately vindicated, however, because of the paucity of practicable alternatives. What critics of the profit motive typically would like to see institutionalized are more *directly* cooperative forms of economic organization (such as the “community of associated producers”). What they object to, in the end, is simply the fact that capitalism uses an indirect strategy to achieve the desired outcome. The challenge for critics of the profit motive, therefore, involves showing that one can actually achieve the end in some other way. This is a burden of proof that has so far proven impossible to discharge.

16.1 WHAT IS PROFIT?

Before getting into the details of this discussion, it is worth taking the time to develop an accurate account of what is meant by “profit” and the sense in which corporations are said to maximize it. My intention here is not to bring in unnecessary technicalities, but rather to get clear on the topic of discussion. Among critics of capitalism, there is a great deal of loose talk about “profit,” with different implicit definitions of the term being used. Furthermore, the current trend in the business world is to regard the concept of profit as somewhat outdated, and instead to describe business corporations as seeking rather to maximize “shareholder value.” Similarly, the search for profit is often described by critics as the primary, or even sole, “objective” of the firm (for example, Bakan 2003), although the reality is significantly more complicated than that.

The first and most important thing to understand about profit is that it is a residuum (Obrinsky 1983). It is sometimes defined as the excess of revenue over expenditure, but this can be misleading, depending upon how expenditure is understood. It is best therefore to think of profit as what is left over after all of the firm’s contractual obligations have been met. This is, I should note, not the Marxist definition, since Marx used the term profit to describe any payment to the providers of capital. In contemporary finance, by contrast, it is standard to distinguish two forms of capital provision: equity and debt. Debt includes both loans made to the firm and bonds that it has sold. The central characteristic

of debt is that its holders are paid an amount that is contractually fixed, what is commonly known as *interest*. Equity, by contrast, also involves the provision of capital to the firm, but on terms that are contractually unspecified. Instead of getting a fixed interest payment, equity investors (that is, shareholders) get a residual claim, which is to say an entitlement to that which is left over after all others—including all bond-holders and creditors, but also workers, suppliers, and so forth—have been paid. This is what we call “profit.” The most obvious feature of equity, in this regard, is that it is much riskier than debt. There is a lot more upside, because in principle shareholders get everything left over, but also a lot more downside, because shareholders are the last to get paid (and in the case of bankruptcy, lose their stake entirely).

A more intuitive way of illustrating the distinction between equity and debt is to compare it to the collaboration that normally takes place in the purchase of a home. There are usually two sources of capital involved in the purchase: the first is supplied by the owner, who is expected to put up a “down payment” representing some fraction of the value of the home, and the second is supplied by a bank, which provides the balance of capital required in the form of a mortgage loan. The owner, in this case, is essentially providing equity, while the bank is supplying debt. This is reflected in the fact that the owner gets the residual claim, which means that if the value of the house goes up, all of that increase in value accrues to the owner, and if the value of the house goes down, the owner suffers the entire loss. The bank, meanwhile, gets no greater payment than the interest rate that was contractually agreed upon. At the same time, it has the right to demand payment, so that if the homeowner fails to meet the terms of the contract, the bank can force sale of the asset in order to recover the sum that it is entitled to. Its position therefore is less risky—it has significantly less upside, but also enjoys significant protection against the downside risks. In effect, in return for getting to keep any increase in the value of the home, the owner also agrees to absorb 100 percent of the losses that the property may suffer; the bank begins to absorb losses only after the owner’s stake has been completely wiped out (in other words, when the foreclosure value of the home becomes less than the outstanding mortgage).

Similarly, the key to the bargain accepted by shareholders in the firm is that they get significant upside to their investment in return for accepting considerable downside risk. What makes the arrangement attractive for the firm is the flexibility that it gets from having a significant amount of equity capital. Firms often experience fluctuations in revenue, and the central limitation of debt is that it is inflexible. Banks demand regular payments of interest on loans and may force the firm to liquidate assets if payment is not forthcoming. Shareholders, on the other hand, need not be paid; if the firm hits a rough patch, the equity serves as a “cushion,” protecting it from the demands of creditors, as well as any other group with a contractual claim on the firm, such as workers. (It is for this reason that shareholders are often also described as “residual risk-bearers.”) Similarly, the firm may, at its discretion, decide to retain profits in order to finance new investment, knowing that shareholders need not be paid right away or on a regular basis. The firm may be punished for such a decision by a lower stock price if investors decide that on balance they are not getting a good enough return, but none of this represents an

existential threat to the firm—unlike with creditors, who can force the firm into bankruptcy if they are not being paid.

One can see then why it is inaccurate, or at least very misleading, to describe the earning of profit as the sole objective of the firm (or to claim that corporations “only care about profit”). In fact, corporations normally assign greater priority to their contractual obligations—such as meeting payroll, making payments to suppliers, and servicing their loans—precisely because they have no discretion in these matters, and failure to make these payments threatens their ability to remain in business. They do, of course, also try to make a profit; it is just that the only legitimate way to do so is by first meeting all of these other obligations. This is, in part, why profit is taken to be a measure of success: it shows that the firm is managing to meet all of its diverse obligations, with “room to spare,” so to speak. Naturally, significant profits will increase the share price, allowing the firm to raise capital by creating and selling new shares, and thereby to expand its business. In the process, it provides a mechanism through which—from a social point of view—investment funds are channeled to those who are likely to make the best use of them.

Also, it should be noted that, although profits that are not reinvested are often disbursed to shareholders in the form of dividend payments, many firms elect not to do so. Indeed, there are many firms that technically have never paid out any “profits” to shareholders at all. They may simply accumulate the funds in order to finance acquisitions, research and development, or other growth strategies. Indeed, many investors buy stock not because they expect it to pay dividends but because they expect the value of the stock to appreciate. For instance, Microsoft Corporation paid out no dividends from 1986, when it first “went public,” until 2003. It is often a symptom of a firm becoming “mature,” “blue chip,” or perhaps “stagnant,” that it begins to pay dividends. It is also common, in jurisdictions where capital gains are taxed more lightly in the hands of individuals than dividends, for firms to buy back their own shares as an alternative to paying dividends. The buyback leads to appreciation in the value of all outstanding shares, allowing the remaining shareholders to take their profits in the form of capital gains (an increase in the value of the shares they own) rather than dividends. Apart from being taxed at a lower rate in many jurisdictions, this also allows the owners to choose when they want to book the profit (by deciding when to sell their shares). Because of the widespread use of such strategies, it is standard now to speak about “shareholder value” rather than “profit,” in recognition of the fact that there are multiple ways in which benefits can be directed to investors.

The idea that corporations are nothing but profit-making machines is often encouraged (particularly by critics of capitalism) through reference to the 1919 Michigan Supreme Court decision, *Dodge v. Ford*, in which the Ford Motor Corporation was ordered to pay out dividends, asserting that “a business corporation is organized and carried on primarily for the profit of the stockholders” (Bakan 2003: 35–7). And yet, as Lynn Stout has argued—in her aptly titled 2008 paper “Why We Should Stop Teaching *Dodge v. Ford*”—this is a serious misrepresentation of the state of corporate law (primarily in America, but also elsewhere). Despite being widely taught and cited by legal theorists, *Dodge v. Ford* has almost never been cited as a legal precedent in any Delaware court decision

(the jurisdiction in which important US corporate law cases are decided). Furthermore, whatever doctrinal value it may have had has been entirely superseded by the “business judgment rule,” which US courts use to grant enormous discretion to firm executives (Easterbrook and Fischel 1991: 90–108). This principle essentially instructs courts to avoid second-guessing management and board decisions in many areas, including the payment of dividends. As Stout puts it,

courts shield directors from liability under the business judgment rule so long as any plausible connection can be made between the directors’ decision and some possible future benefit, however intangible and unlikely, to shareholders. If the directors lack the imagination to offer such a ‘long-run’ rationalization for their decision, courts will invent one. (2008: 170–1)

This is the rubric, for instance, under which charitable and political donations by corporations are protected from shareholder lawsuits. Corporations may also accumulate enormous piles of cash (“retained earnings”), which can be used, *inter alia*, to finance mergers and acquisitions. Many of these uses of corporate funds are of extremely dubious value to shareholders, but it is difficult to imagine circumstances under which courts would intervene and force the money to be disbursed to shareholders.

It is also worth noting that the maximization of shareholder value, or even the intention to seek profits, is seldom mentioned explicitly in corporate charters. There is of course an implicit understanding that this is what the firm intends to do, but there is no explicit contractual commitment, and thus, no legally enforceable obligation. There is also nothing to prevent businesses from assigning other objectives priority in their charters, and there are a number of celebrated examples of corporations that do. Perhaps the most high-profile is the *New York Times* Corporation, which assigns highest priority to its news-reporting functions, and secondary importance to profits (Easterbrook and Fischel 1991: 13). This is why certain progressive corporate law scholars have opposed the development of so-called “benefit corporations” in the United States, which are investor-owned, but have an explicit commitment to certain charitable objectives, as well as to consideration of the interests of non-shareholder groups. The central argument against them is that they are redundant, because there is nothing to prevent the organization of a standard business corporation in the same way.

As Frank Easterbrook and Daniel Fischel put it,

If a corporation is started with a promise to pay half of the profits to the employees rather than the equity investors, that... is simply a term of the contract. It will be an experiment... Similarly if a bank is formed with a declared purpose of giving priority to loans to minority-owned businesses or third-world nations, that is a matter for the venturers to settle among themselves. So too if a corporation, on building a plant, undertakes never to leave the community. (1991: 36)

This is, of course, rather tongue-in-cheek, because it would be difficult for a standard business corporation to attract investors if it were to announce straightaway that it is not

going to be seeking to maximize shareholder value. The essential point though is an important one: the firm is established through contractual relations between the involved parties, and corporate law grants them extremely wide discretion when it comes to deciding the terms of these contracts. The image of a legal straitjacket, forcing firms to pump out profits for investors, is entirely illusory.

Finally, it may seem rather obvious, but it nevertheless bears repeating, that the slippage one often encounters in popular rhetoric between “profit” and “self-interest” or “greed,” is extremely misleading.¹ Adam Smith may have used the two terms interchangeably when discussing the “interest” of the butcher or the baker, but that is because he was writing at a time before the emergence of the modern business corporation. With the development of professional management, the people who are making the profits and those who are receiving them are usually two different groups. Indeed, many firms go to great trouble in attempting to “align” the incentives of the two by giving managers an equity stake in the firm, such as stock options in lieu of salary. Nevertheless, a major structural feature of the modern business corporation is that senior executives have discovered various ways of enriching themselves at the expense of shareholders (Bebchuk and Fried 2009). For example, the cost of stock options is borne almost entirely by shareholders, in the form of share dilution. Thus all the loose talk about corporate “rapaciousness” or “greed” is in many ways quite unhelpful because it obscures the complex mixture of incentives that structures the actions of the various parties involved in the firm. Also, it fails to take seriously the fact that managers often do not stand to gain personally from the various forms of corporate misconduct that they suborn. Even talk of the “profit motive” is potentially confusing because it blurs the distinction between profit as an *organizational objective* and the specific incentives that individuals within the firm are given (such as salary, bonuses, and equity-based compensation). This becomes extremely important when it comes to thinking about alternatives to the profit maximizing firm because, even when it is possible to eliminate profit as an organizational objective, this does not automatically eliminate the incentives that may have been pushing managers to engage in various forms of malfeasance.

16.2 WHY PROFIT-SEEKING IS COLLECTIVELY SELF-DEFEATING

There is an old saying that every free market is a failed cartel. Firms recognize that it is not in their interest to compete with one another, and so left to their own devices they will often find ways to avoid it. In some cases, this will take the form of an explicit cartel, with rules and regulations limiting competition, official meetings to set prices and

¹ For example: “The corporation’s legally defined mandate is to pursue, relentlessly and without exceptions, its own self-interest, regardless of the often harmful consequences it might cause to others” (Bakan 2003: 1–2). For analysis of the issue, see Flew (1976).

divide up territory, and all the other trappings of a formal organization. States generally prohibit this sort of collusion. Furthermore, in cases of natural monopoly—in which it is simply too inefficient to have multiple, competing firms—states will normally either nationalize the sector or else regulate the prices that can be charged, effectively limiting the profits that can be earned. Thus there is no general license given to firms enabling them to pursue profits however and wherever they see fit. They are given specific license in particular sectors or areas in which it is possible to ensure a reasonable degree of competition between multiple firms (or a plausible threat of entry from new firms).

When it is possible to organize a competitive market, the pursuit of profit is important because it is what drives the competition. It is precisely because the owners of the firm are able to appropriate the entire residuum that there is such a powerful incentive—transmitted more or less effectively to managers—to chase windfall profits. It is important to recognize, however, that when firms do so, they are in effect adopting a free-rider strategy vis-a-vis other firms in the market. In other words, for the firms involved, competition has the structure of a prisoner’s dilemma. This feature of markets is, unfortunately, obscured in the standard textbook treatment, which focuses on perfectly competitive markets, where firms have no influence on prices and only vary their supply decisions. It helps to clarify the structure of real-world competition if one begins with a simple strategic model of competition between two firms.

Consider, for instance, a set of supply and demand curves such as shown in Figure 16.1 (drawn as straight lines for convenience). One can see why \$4 is referred to as the “market-clearing” price: it is the price level at which the total quantity of the good demanded by the consumer is equal to the total quantity that the producer is willing to sell (point *b*). At a higher price, such as \$6, the consumer is only willing to purchase two units (point *a*)

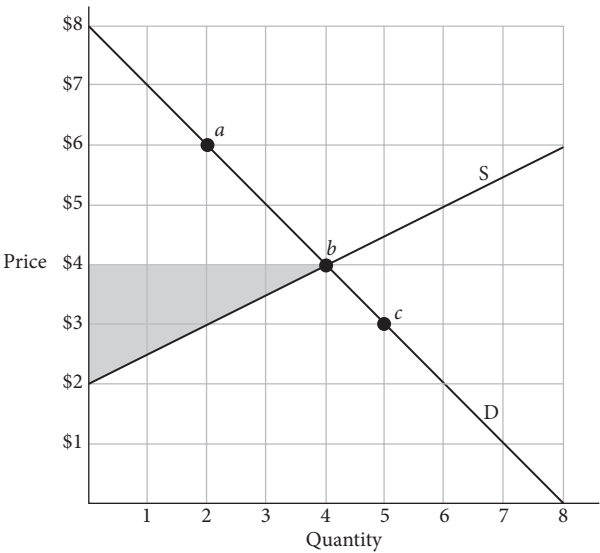


FIGURE 16.1 Supply and demand.

whereas the producer would be willing to sell eight units. At a lower price, such as \$3, the consumer would be willing to purchase five units whereas the producer is only willing to sell two units (point *c*). The fact that the supply curve rises is a reflection of the fact that the producer has variable costs, which are assumed to increase as the number of units produced increases. The producer's surplus is defined as the revenue earned from a sale minus these variable costs. In Figure 16.1, it is represented by the shaded area between the price line and the supply curve. If the price is \$4, and there is only one producer, who sells four units, the producer's surplus will be \$4. If the firm has fixed costs of \$4, this results in a net profit of zero.

Suppose that there is only one firm in the market. This means that it is able to exercise "market power," or the ability to influence prices by changing its supply decisions. For instance, if it offers only two units for sale, and is able to credibly commit to selling no more (for example, by making a "take it or leave it" offer), then it will be able to sell two units at a price of \$6. As a result, its producer surplus will increase to \$7, resulting in a net profit of \$3 (as shown in Figure 16.2). This outcome, it should be noted, is also inefficient. At a price of \$4, the sum total of producer and consumer surplus is \$12, with \$8 (or its utility equivalent) going to the consumer and \$4 to the producer. Raising the price to \$6 and reducing the quantity transacted to two units reduces the total gains from trade to \$9, implying a "deadweight loss" of \$3. This is advantageous to the producer, however: out of the diminished surplus, the producer gets \$7 and the consumer is left with only \$2. One can see why, in this case, profit-seeking is anti-social, because the firm's profits are earned entirely at the expense of the consumer. One can also see why, as a matter of social policy, one might want to see trade occurring at market-clearing prices, because the latter maximize the benefits of the exchange.

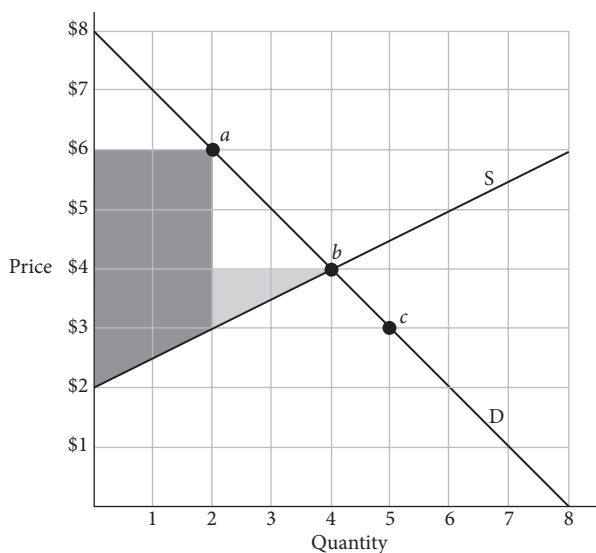


FIGURE 16.2 Exercising market power.

One way of achieving this policy objective is to introduce multiple buyers and sellers into the trading relation. The ability of the producer to charge more than the market-clearing price depends upon the ability to make a credible “take it or leave it” offer. But if more producers are introduced, they may begin to compete with one another for sales. When they do this, each one *loses the ability to hold out*. This is the primary benefit of marketplace competition. If one seller tries to hold out for a high price, the other sellers will underbid him in order to unburden themselves of excess merchandise. Similarly, if one buyer tries to hold out for a low price, the other buyers will outbid her. As a result, the only equilibrium price level in a competitive market is the market-clearing price, which also happens to be the only price that generates the efficient allocation of goods. This is the key to the success of markets: trade provides a source of Pareto improvements, and competition provides the dynamic that drives economic agents to exhaust all of these potential improvements.

Suppose, for instance, that there are two producers, who start offering their goods at a price of \$6. They will only get offers to purchase two units of the good. Since the producers are willing to sell eight units at this price, there is an excess of supply over demand; this means that they will both be left with unsold goods. The big question then becomes *who* will get stuck with these unsold goods (or, equivalently, who will get what portion of the producer’s surplus). This presents each individual producer with an opportunity. Rather than run the risk of getting stuck with unsold goods, the firm could just drop its asking price slightly lower than its competitor, thereby guaranteeing that all of its goods will sell. The losses that are suffered from lowering the asking price will be made up for by the larger portion of the producer’s surplus that can be captured. However, since the goal is to select an asking price lower than one’s competitors, and all producers are in the same situation, the interaction has the structure of a race to the bottom.

Consider the following scenario for a competition between two producers: if the producers have the same asking price, the consumer will purchase an equal amount from both of them. If one of the producers has a lower price than the other, the consumer will purchase as much as he can at the lower price, moving on the higher-priced producer only if the lower-priced producer runs out. So if the producers start out with an asking price of \$6, the consumer will buy only two units, one from each supplier. This might tempt one producer to drop the asking price to \$5 in order to sell more units. The choice that each one faces is shown in Figure 16.3. The key to understanding the dynamics of this interaction is to pay attention to the incentives that each individual supplier faces.

The only equilibrium of this game is for each producer to charge \$5. The reasoning behind this is straightforward. No matter what the other producer does, each can do better charging \$5. If the other asks for \$6, then one can split the \$3 profit by charging \$6 as well, or lower one’s price to \$5 and steal all of the business away, thereby selling all three units and capturing the entire profit of \$2.66. Of course, lowering the price reduces the size of the profit, but the fact that the firm gets all of it, rather than half of it, is enough to outweigh the loss. On the other hand, if the other firm is going to charge \$5, then the size of the surplus is going to be smaller regardless of what one does, and so it is best to ensure that one gets one’s share of the market. As a result, it is also in one’s interest to charge \$5. So either way, the firm should reduce its price to \$5.

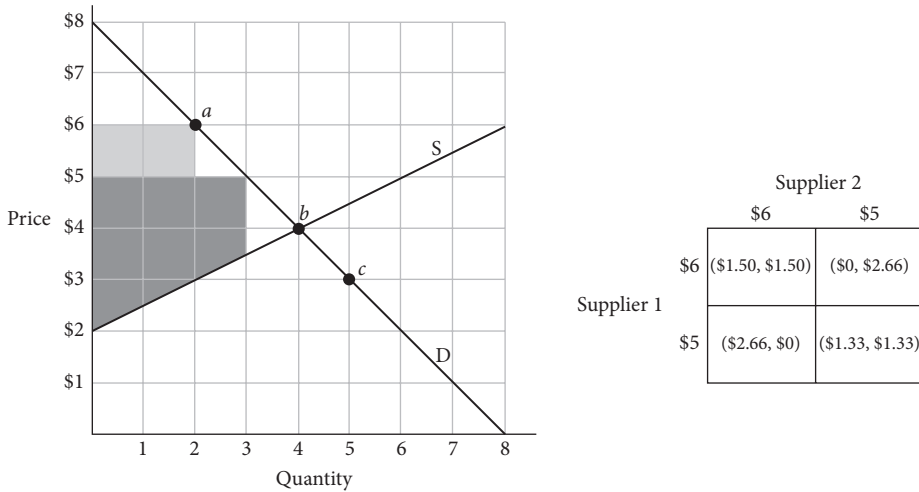


FIGURE 16.3 Competition among suppliers.

Thus the two suppliers are stuck in a classic prisoner's dilemma. High prices generate a relative "scarcity" of customers by reducing demand below the level of voluntary supply, which in turn encourages suppliers to compete for customers by offering lower prices. Thus both will reduce their price from \$6 to \$5, despite the fact that this reduces the total size of the suppliers' surplus. This competition continues until customers are no longer scarce, and neither firm has unsold goods, which is, of course, the point at which supply and demand curves intersect. (Thus the suppliers are not just stuck in a prisoner's dilemma, but also a race to the bottom.)

This example shows how, when prices start above the market-clearing level, competition between sellers has the effect of driving these prices down. The situation that arises when the price starts out below the market-clearing level precisely mirrors the one in which it starts out above, except that this time it is consumers who begin to compete with one another instead of suppliers. At \$3 per unit, consumers would like to buy five units of the good, while suppliers are only willing to sell two. Thus the good will become scarce, which is also to say that there will be "pent-up" demand: not everyone who wants to buy the good will be able to get some of it. The big question then becomes *who* will get left out (or equivalently, who gets what portion of the consumer's surplus). This presents each consumer with an opportunity. Rather than run the risk of getting stuck with nothing, she could offer the suppliers slightly more than the going price. The losses that she suffers from paying more will be outweighed by the fact that suppliers will not run out of the good before she gets some. This competition among purchasers drives prices up toward the market-clearing level. At that point, all of the consumers get as much as they want, so there is no reason for them to compete with each other any further.

Incidentally, the competition among sellers and among buyers can take a variety of different forms. For example, the competitive dynamic can be seen in the way that "sales" are managed in the retail sector. Many retailers start out pricing things a bit high,

then put them on sale when it looks as if they will not all sell. The kind of competitive price-cutting that may result is quite visible. Competition among consumers is considerably less visible. Consumers effectively compete with one another by buying things *before* they go on sale. That is, many consumers will not pay full price for a good if they expect it to go on sale in a month or two. Naturally, if everyone waited, then things would go on sale. But if people think the seller will run out of the good, they will buy it at full price rather than risk waiting.

This example shows clearly how both suppliers and consumers face a collective action problem. Each consumer who holds off on a purchase in order to wait for a sale produces a positive externality for other consumers in the form of increased pressure on suppliers to lower the price. Similarly, each supplier who delays putting things on sale produces a positive externality for other suppliers in the form of increased pressure on consumers to buy at that price. In both cases, this generates a free-rider incentive: consumers may “break ranks” and buy at full price, or suppliers may “break ranks” and have a sale. The consequence of these two collective action problems will be downward pressure on the price of plentiful goods and upward pressure on the price of scarce goods. The only equilibrium will be the point at which the amount of the good exchanged is just right. Markets will clear, and the resulting (partial) allocation will be Pareto-optimal.

Finally, it is worth noting one peculiarity of this analysis, which is that it represents marketplace competition as occurring only amongst producers and amongst consumers. It does not, however, have producers adopting a competitive stance toward consumers, or vice versa. Yet there is a more commonsense way of thinking about how market interactions occur that sees the central antagonism, as well as the focus of profit-maximization, involving relations between buyers and sellers across the price line. One way of representing this competition, following John Kenneth Galbraith (1952), would be to assume that parties on both sides are attempting to develop market power, perhaps through tacit collusion, in order to extract rent from the opposing side. The central “competition” in such markets then involves the development of “countervailing power,” typically on the part of consumers, in order to block the rent-extraction that producers have been able to achieve. The role that retailers play in securing bulk discounts from producers is an example of this phenomenon (essentially acting as agents of the consumers). This back-and-forth over rents has a similar effect to the more conventional forms of supply-side and demand-side competition in that it pushes prices and volumes transacted closer to market-clearing levels.

16.3 MARKET ADVERSARIALISM

What I have described in the previous section are a number of very well-known features of market economies. Competitive buying and selling function together as a revelation mechanism, taking an invisible set of consumer preferences and a not-easily-observable set of production possibilities and transforming them into a publicly observable piece of

information—a price—that reflects the *scarcity* of each good relative to those preferences and possibilities. Trying to get along without this information in economic decision-making is a recipe for colossal waste, whereas the only other ways of getting at it are demonstrably inferior (Kornai 1992; Roemer 1994). So there is not much room left for controversy over the merits of competitive markets (at least not at the current level of technological development), and because it is the profit orientation of firms that generates the competition, these observations would appear sufficient to settle the question about the moral status of profit.

For those who hold a consequentialist view of morality, all of this is entirely unproblematic. The narrow profit orientation of firms may at first glance seem anti-social, but if one looks at its broader consequences, the beneficial effects achieved through the operation of the price system can easily serve to justify that orientation. Marketplace behavior therefore does not pose any special difficulties; it merely has a connection between action and consequence that is slightly more complicated than usual. Yet, the ease with which the various “sharp practices” of the marketplace can be justified might also be held up as an example of what many consider to be the major *problem* with consequentialist views of morality: they make it entirely too easy to justify any sort of behavior, because the exclusive focus on consequences obliterates a variety of distinctions that are important for assessing the moral quality of actions. This is why, for instance, consequentialists have had little to contribute to just war theory; the view makes it too easy to justify too many things, including many acts that are currently considered war crimes. Consistent consequentialism would be radically revisionary with respect to everyday morality, opening the door not just to marketplace competition, but to various forms of ruthless, disloyal, unprincipled behavior. Thus consequentialism fails to articulate what many people find so problematic about marketplace interactions, which is the tension that they experience between the norms of behavior in business and in other areas of life.

Regardless of the stance one takes on the overall merits of consequentialism, it does seem clear that the profit orientation of firms does not pose any particular problem for it (Applbaum 1999: 177). It is going to be a problem, or a serious problem, only for deontological views of morality, or views that take the constraints of everyday morality more seriously, because so much of everyday morality is oriented toward encouraging individuals to act more cooperatively.² The standard rules against lying, cheating, and stealing are all intended to prohibit individuals from adopting what amount to free-rider strategies in interactions that have the structure of a collective action problem. More abstract formulations, such as the “golden rule,” which instructs us to treat others in the way that we ourselves would like to be treated, have a similar effect. Finally, Kant’s first formulation of the categorical imperative, which requires individuals to act only on principles that they could will to become a universal law, prohibits free-riding in any

² This includes various forms of virtue theory that adopt a strict definition of vices such as avarice, cupidity, or even just dishonesty, such as MacIntyre (1984). In this chapter, however, I will focus only on deontological theories.

form; at one point, Kant glosses the principle as one that prohibits “making an exception” for oneself (1785: 424). All of these principles would appear to prohibit the sort of competitive behavior that is expected of market actors. The ambition of the free-rider, after all, is to take advantage of others who are *not* free riding. And if everyone is free-riding, then the situation is one in which everyone is voluntarily choosing an outcome that is worse for everyone. Either way, the intention cannot be universalized or the outcome willed universally, and pointing to the beneficial by-products of the collective action problem is, from any of these perspectives, insufficient to overcome the deontic prohibition.

Thus, moral theories that focus on reciprocity or universalization of motive generally have difficulty tolerating competitive interactions, simply because these are structured in such a way as to, as Waheed Hussain puts it, “pit people against each other” (2018). Furthermore, they seem to instrumentalize social relations in a way that is at least *prima facie* objectionable, because competitors treat each other in ways that are not intrinsically justifiable, or justifiable with respect to either agent’s own intention, but can only be justified through reference to some ulterior purpose or effect of the interaction. This is a consequence of the fact that markets are an *adversarial* institution, which is to say one whose social function is discharged, not by having participating individuals intend the desired outcome, but rather by having them adopt an impossible set of intentions and letting the desired outcome arise as a (typically unintended) consequence of the agonistic interaction that results (Martin 2015). In this respect, business ethics turns out to be a species of what Applbaum (1999) refers to as “adversarial ethics.”

The central challenge in the field of adversarial ethics is two-fold. It involves developing a framework that can explain how it could be permissible for agents to adopt an “adversarial stance” that licenses them to act in ways that violate certain deontic prohibitions, such as acting on the basis of non-universalizable maxims or ignoring a certain class of reasons for action. But it also involves articulating the limits of the adversarial stance, such that it stops short of offering *carte blanche* to engage in any sort of antisocial behavior. In other words, the central challenge is to develop a normative reconstruction of the *scope* and *limits* of the adversarial stance.

Within the field of adversarial ethics, there are some widely accepted principles that have been identified that serve to reconcile the adoption of an adversarial stance with the more general impartiality constraints of deontological morality (Applbaum 1999: 4). In competitive sports, for instance, the fact that participation in a game is entirely voluntary is considered sufficient to lift many of the usual constraints on physical interference and assault that structure everyday interactions. In other cases, the mere fact that the nature of the interaction is common knowledge among participants is sufficient to loosen the ordinary constraints. The job of the opposition party in a parliament, for instance, is to oppose the policies of the government. Thus it would seem unreasonable to accuse members of “lying” when their public pronouncements happen not to reflect their personal convictions; people know how the game is played and are thus highly unlikely to be misled by the utterance.

Neither of these two options, however, is particularly persuasive in the case of marketplace interactions. Most obviously, participation in the market economy cannot

be considered voluntary, given the very high cost of withdrawal, combined with the fact that the private property regime is coercively imposed. This is the major reason that business ethicists have long condemned the analogy to sport. Similarly, the mere fact that everyone knows “how the game is played” provides little solace. With the example of politics, the “lie” turns out not to be much of a lie, because common knowledge of the adversarial structure of the interaction allows everyone to interpret the utterance correctly (that is, not in accordance with its literal meaning). In the case of marketplace competition, by contrast, competitive actions have negative consequences for others that are in no way attenuated by common knowledge of the structure of the interaction.

Applbaum’s own suggestion is that the adversarial behavior associated with a particular institutional role may be justified if it is “necessary” to the discharge of the role and if the institution as a whole is one that can be unanimously or universally endorsed (perhaps on the basis of its good consequences). Thus the suggestion is often made that, when it comes to marketplace interactions, an adversarial action is justified if it “leads to good consequences in equilibrium” and if the action “is necessary to produce such good consequences in equilibrium” (1999: 188). This approach has its merits, but the requirement that the *action* be necessary is overly restrictive when it comes to market competition. For instance, it will often be the case that if one firm does not take advantage of a particular opportunity, another firm will come along and do so, as a result of which there is no necessary connection between any one firm’s actions and the good consequences that result. More generally, because market institutions are aimed at maximization—in this case, of the gains from trade—one would not want to say that each firm is *obliged* to perform the one action that will produce the very best outcome, and anything short of that is forbidden.³ Thus Applbaum also formulates a rule-based version of the constraint, which specifies that an adversary action may be justified if “the rules permitting such action are necessary for the continued success or stability of the game as a mutually advantageous cooperative venture” (1999: 123). As he continues: “If a rule allowing sharp adversary tactics is necessary for the success of a cooperative scheme, and the general practice of regulation-by-rule is necessary, each adversary act under the rule need not be necessary for the act to be permitted” (126).

On this view, markets generate a broad set of permissions for firms to engage in a variety of forms of competitive behavior without really requiring any particular strategy. The pressure toward maximization is generated by the structure of the collective action problem and therefore need not be imposed as a moral obligation. Thinking about why societies adopt adversarial institutions in particular domains may help us to understand why they have this structure. The first step involves recognizing that, perhaps with the

³ Waheed Hussain makes a similar point, objecting to the way that a welfare-consequentialist justification of marketplace behavior suggests that “corporations are normatively required to maximize their long-run market value by maximizing profits” (2012: 318). His solution, however, involves introducing a “personal sphere” that is insulated from this maximization imperative. In my view, the problem arises from the relevant “normative requirement” being interpreted as an obligation (which does follow from the consequentialist perspective, as well as certain versions of Applbaum’s). The more natural way of resolving the difficulty is simply to interpret the requirement as a permission.

exception of sport, adversarial institutions exist only in the realm of non-ideal theory. In an ideal world, in which everyone always did what they were supposed to do, there would be no need for adversarial institutions. If one considers the desired outcome of such an institution, the first-best arrangement is typically to have that outcome achieved directly by imposing a moral obligation on everyone to intend its realization. It is typically because of *compliance problems* with this first-best arrangement that direct normative prescription is abandoned in favor of the indirect approach achieved through adversarialism (Heath 2014: 181).

Consider, for instance, a criminal trial procedure. Ideally, people would just stand in front of the judge and tell the truth. We recognize, however, that they usually have an incentive to lie—indeed, the court itself often generates those incentives—and as a result, one cannot expect the norm of honesty to be respected. So instead of insisting on that norm, and using whatever forms of moral suasion are available to achieve its approximate realization, we instead adopt an adversarial procedure by instituting a competitive testing of claims. We release the accused from the obligation to present a fair, balanced, or unbiased account, and instead grant permission to do what one would expect the accused to do anyhow, which is to present a self-serving and self-exculpatory account. Then, in opposition to this, we set up a prosecutor, whose job is to present an equally partial, but in this case damning, account. The thought is that by comparing these two accounts, both of which are maximally distorted but in opposite directions, the judge will be in the best position to determine the truth of the matter.

The situation with respect to the economy is not all that different. Ideally, one could have an entirely cooperative system of production, but every attempt to institutionalize such an arrangement has been undone by significant compliance problems. Under Soviet-style central planning, for instance, factory managers were expected to state honestly what inputs would be required in order to meet certain production quotas. The incentive, however, was to overstate these requirements as much as possible, not just to decrease the chances of missing the target, but also so that surplus inputs could be stockpiled to be later sold on black markets or bartered (Kornai 1992: 268; Rutland 1993: 76–7). This was often done without any motive of personal graft or corruption, but simply as insurance against unexpected events that might result in a missed quota, such as a broken machine or a shortfall in some other input. Having a large stockpile of raw materials allowed managers to obtain other goods or services that, despite being required by the production process, could not be obtained through official channels under the constraints of the current plan. The net effect, however, was that vast quantities of raw materials and intermediate goods wound up sitting idle because plant managers preferred to hold them back rather than feed them into the actual production process.

The willingness to accept private ownership of the means of production rests, in part, in the recognition that it is simply too much to expect managers to care about the overall consequences for society of these sorts of productive inefficiencies. The first step, however, is the counterintuitive one of granting private individuals permission to appropriate whatever surpluses can be achieved in the production process. It is the second step, the introduction of a competition between multiple firms with market pricing of both

inputs and outputs, that constrains private behavior and makes it so that private-sector managers are forced to act in the way that, ideally, “managers of production in a socialist system” would act (Lange 1937: 123). Private ownership, for instance, eliminates the incentive to stockpile idle resources, because profit is calculated in terms of return-on-capital, and firms must compete with one another for investment. Thus through competition the system achieves a rough approximation of what a socialist system aspires to achieve through direct normative prescription.

One can see from these examples why adversarial institutions, rather than imposing specific moral obligations, instead tend to operate through “deontic weakening” of everyday moral obligations, taking actions that were forbidden and making them permissible: the institutions themselves are typically a second-best response to a set of compliance problems raised by a system of direct normative prescription. It would not make any sense, under these circumstances, to replace the latter system with another, directly prescriptive institution. This would merely be to trade one compliance problem for another. The trick involved, in an adversarial institution, is to take the incentives that were generating the compliance problem, grant individuals permission to act upon them, and impose an external structure that neutralizes the anti-social impact of these actions.

Thus, the right way to think about adversarial institutions is not that the goal of the institution as a whole translates into an obligation to perform the one action that is best, as far as promotion of that goal is concerned, but that the institution provides an exemption from certain everyday moral constraints, which in turn creates a zone of discretion, as a result of which individuals need not intend the goal. This does not, however, make the goal of the institution irrelevant. Although it no longer directly prescribes conduct, it still plays the important role of *constraining the set of permissions*. This is for the obvious reason that the institution cannot dissolve all the bonds of conventional morality; it merely offers selective exemption from particular constraints, sufficient to implement the indirect strategy that it uses to achieve the desired outcome. This is illustrated in Figure 16.4.

Elsewhere, I have argued that the best way of specifying the constraint imposed by the institutional goal is to say that while individuals need not intend that goal, their intentions, and the actions that they choose, must be at least consistent with it in the given institutional context (Heath 2007). In competitive interactions, for instance, this

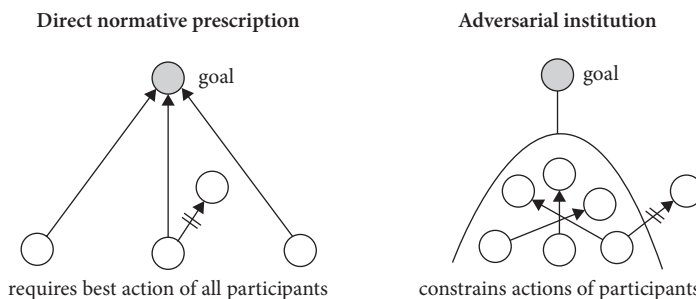


FIGURE 16.4 Direct normative prescription vs. adversarialism.

means that participants need not advance the general goal of the competition with their every action, but they must at least refrain from acting in ways that undermine it. Applying this reasoning to marketplace competition can provide us with guidance on the question of firm behavior. The point of the competition is to achieve a set of market-clearing prices, and thus a more efficient allocation of productive resources, as well as consumption goods, across the economy.

This generates three “families” of permissible competitive strategies:

1. *Supply decisions.* In the classic model of perfect competition, firms are modeled as “price takers,” who have no choice but to vary their supply decisions in response to changes in market prices. This is the most basic permission that firms enjoy: the right to increase or decrease the amount that they produce based on their own calculations of profit and loss. Along with this comes the discretion to vary their purchases of inputs (including labor) as a function of their desired level of output.

2. *Pricing decisions.* Given that firms seldom find themselves in circumstances that resemble perfect competition, they often enjoy some discretion in the prices that they charge for their outputs. This is, in fact, the central mechanism of marketplace competition, as described earlier. Thus firms enjoy a permission to negotiate prices, on both inputs and outputs, without concerning themselves with the interests of the other party or worrying about whether the consequent division of the cooperative surplus is equitable. (This is, it should be noted, a controversial permission, since the right to charge whatever price they think the market will bear often results in firms raising prices in circumstances where the need is greatest.)

3. *Shifting the supply curve down.* The “supply curve” in a diagram such as Figure 16.2 typically represents an aggregate of the supply curves of several different firms. Under real-world conditions, there will actually be considerable variation among firms with respect to their individual supply curves. As a result, the distribution of producer’s surplus will be unequal among these firms. A firm that is able to shift its own supply curve down, relative to other firms in the sector, will be able to capture a greater portion of this surplus. There is a vast set of competitive strategies that constitute permissible ways of doing this, the most obvious being increased efficiency in production achieved through cost-reduction, process improvements, technological innovation, and so forth.

Within each of these sets of strategies, there are of course some that will be impermissible as well. The firm’s supply curve can be shifted down by externalizing production costs rather than actually lowering them. Both pricing and supply decisions can be aimed at exploiting market power, resulting in deadweight losses. Sorting out one from the other is the “heavy lifting” involved in business ethics. It is, however, not a task for philosophers, but rather for managers, those who confront the day-to-day task of running a successful business while maintaining their moral integrity. The need for philosophical reflection arises only because everyday morality is unreliable as a guide to what is permissible and what is impermissible in this domain. Thus the task of the philosophers is to articulate the higher-level goal, along with the broad set of constraints in terms of which particular competitive strategies must be assessed.

16.4 CONCLUSION

G.A. Cohen used to tell the story of his father Morrie, who began working as a dress-cutter in a garment factory at age 14 and retired 55 years later, “still working in a factory as a dress cutter” (2000: 180). The ending, however, was not so happy:

One December day in 1979 the boss of the factory in which he was working called three dress cutters into his office: Morrie, and two younger ones. The boss told them that there was not enough work to keep them on, and that he would therefore have to let them go . . . Morrie was dismissed because it no longer paid the boss to pay him . . . Business is, among other things, people treating people according to a market norm—the norm that says they are to be dispensed with if they cannot produce at a rate which satisfies market demand. Of course that promotes “efficiency,” but it also corrupts humanity. Business turns human producers into commodities. (180–1)

One can see in this story many of the characteristic features of the moral critique of capitalism. First there is a vagueness of the complaint: “it no longer paid the boss to pay him.” Because Cohen was a man capable of great analytic precision, one must assume that this is intentional. Partly it is because he understood what the counterarguments were, such as that it is inefficient to keep on workers when there is insufficient demand for the products of their labor. Yet Cohen obviously could not escape the sense that there is something *wrong* with the market arrangement. What rankles is the fact that Morrie was cast aside, despite decades of loyal service, whereas one would think that a long period of employment would generate some entitlement or reciprocal loyalty from the employer. And yet the “market norm” denies this, and indeed, voids whatever obligations the employer might otherwise have had.

My own view is that Cohen’s moral intuition is a genuine one, but that it is nevertheless permissible for Morrie’s “boss” to fire him, because markets provide him with an exemption from the everyday moral constraints that Cohen is tacitly appealing to. Indeed, the central problem with a consequentialist analysis is that it is unable to do justice to the initial intuition. From a consequentialist perspective, the problem with Cohen’s story is that he is thinking too much about his father’s welfare while ignoring the benefits that flow from Morrie’s dismissal, such as the increased happiness of consumers able to purchase dresses at lower prices and the increased returns to shareholders of the firm. For the consequentialist, the “moral intuition” at the heart of the story is really just a case of flawed moral mathematics.

The central virtue of the adversarial analysis is that it is less dismissive, and therefore more capable of articulating the force of Cohen’s moral intuition. Normally, if one was involved in a cooperative relationship with someone like Morrie over the course of decades one would owe it to him to remain committed until he chose to retire, or some mutually agreeable way of winding down the relationship could be established. Business, however, permits unilateral termination. What adversarial ethics tries to explain is why

business constitutes a special circumstance, which grants individuals exemption from some of the everyday norms of reciprocity that would prohibit such termination.

Of course, it would be going too far to say that there is an *obligation* on the part of the firm to fire Morrie. It is in fact another downside of the consequentialist analysis that it can easily generate this implication (Hussain 2012). But it is also an implication of any analysis of adversarial institutions that licenses only those actions that are necessary to achieve some institutional goal. My goal here has been to explain why the adversarial institution of the market generates only a permission to fire Morrie. There is usually a certain amount of slack within business organizations, and a great deal of cross-subsidization between units, as well as between employees (Frank 1986). So if the firm had wanted to keep Morrie on, it probably could have done so. The important point is simply that, pace Cohen, it was not *bound* to do so. Indeed, it was permissible for the firm to fire Morrie, on the basis of no more than the proximate objective of increasing its profits.

This *sounds wrong*, which is why the moral critique of capitalism in general, and the profit motive in particular, remains so persistent, and why it is rediscovered—as though it were a fresh insight—by each new generation. What is needed is a normative theory that it is able to grant the force of the moral intuition without in the end succumbing to it. The analysis developed here is intended to explain why the norms that govern marketplace behavior, including first and foremost the imperative of profit-maximization, do not correspond to the way that *in general* we want people to treat one another, but that *in this particular circumstance* they are justifiable.

REFERENCES

- Applbaum, Arthur. 1999. *Ethics for Adversaries*. Princeton, NJ: Princeton University Press.
- Bakan, Joel. 2003. *The Corporation*. Toronto: Penguin.
- Bebchuk, Lucian, and Jesse Fried. 2009. *Pay Without Performance*. Cambridge, MA: Harvard University Press.
- Boltanski, Luc, and Eve Chiapello. 2005. *The New Spirit of Capitalism*. G. Elliott (trans.). London: Verso.
- Cohen, G. A. 2000. *If You're an Egalitarian How Come You're So Rich?* Cambridge, MA: Harvard University Press.
- Cohen, G. A. 2009. *Why Not Socialism?* Princeton, NJ: Princeton University Press.
- Easterbrook, Frank H., and Daniel R Fischel. 1991. *The Economic Structure of Corporate Law*. Cambridge, MA: Harvard University Press.
- Flew, Anthony. 1976. "The Profit Motive." *Ethics* 86: 312–22.
- Frank, Robert. 1986. *Choosing the Right Pond*. Oxford: Oxford University Press.
- Galbraith, John Kenneth. 1952. *American Capitalism*. New York: Houghton Mifflin.
- Giroux, Henry. 2008. *Against the Terror of Neoliberalism*. New York: Routledge.
- Greenfield, Kent. 2006. *The Failure of Corporate Law*. Chicago, IL: University of Chicago Press.
- Heath, Joseph. 2007. "An Adversarial Ethic for Business: or, When Sun-Tzu Met the Stakeholder." *Journal of Business Ethics* 72: 359–74.
- Heath, Joseph. 2014. *Morality, Competition and the Firm*. New York: Oxford University Press.
- Hussain, Waheed. 2012. "Corporations, Profit Maximization and the Personal Sphere." *Economics and Philosophy* 28: 311–31.

- Hussain, Waheed. 2018. "Pitting People Against Each Other." Manuscript.
- Kant, Immanuel. 1785. *Groundwork of the Metaphysics of Morals*. Mary Gregor and Jens Timmermann (trans.). Cambridge: Cambridge University Press (2012 edition).
- Kornai, János. 1992. *The Socialist System*. Oxford: Clarendon.
- Lange, Oskar. 1937. "On the Economic Theory of Socialism, Part 2." *Review of Economic Studies* 4: 123–42.
- Lipsey, Richard G., Douglas D. Purvis, and Peter O. Steiner. 1988. *Economics*, 6th edn. New York: Harper & Row.
- Macintyre, Alisdair. 1984. *Marxism and Christianity*. Notre Dame, IN: University of Notre Dame Press.
- Martin, Dominic. 2015. "There is No Bathing in the River Styx: Rule Manipulation, Performance Downplaying and Adversarial Schemes." *Ethical Theory and Moral Practice* 16: 129–45.
- Obrinsky, Mark. 1983. *Profit Theory and Capitalism*. Philadelphia, PA: University of Pennsylvania Press.
- Roemer, John. 1994. *A Future for Socialism*. Cambridge, MA: Harvard University Press.
- Rutland, Peter. 1993. *The Politics of Economic Stagnation in the Soviet Union*. Cambridge: Cambridge University Press.
- Skidelsy, Robert, and Edward Skidelsky. 2012. *How Much Is Enough?* New York: Other Press.
- Stout, Lynn A. 2008. "Why We Should Stop Teaching *Dodge v. Ford*." *Virginia Law and Business Review* 3: 164–76.
- Walsh, Adrian, and Tony Lynch. 2008. *The Morality of Money*. New York: Palgrave Macmillan.

CHAPTER 17

THE ETHICS OF MONEY AND FINANCE

JOAKIM SANDBERG

17.1 INTRODUCTION

THROUGHOUT most of our cultural history, activities that involve money or finance have been the subject of intense moral scrutiny and ethical debate. It seems fair to say that most traditional ethicists, such as religious and Medieval scholars, held a very negative attitude towards such activities. Think, for example, of Jesus' cleansing of the temple from moneylenders, the condemnation of money as "the root of all evil," and similar expressions in many other older texts. Attitudes in this regard seem to have softened over time, especially in the modern age where the great gains in wealth and welfare due (in large part) to financial activities are evident to most. Nowadays, a great deal of young people are attracted by the high status and income associated with working in "the City" (of London) or on Wall Street (of New York). And their parents do not typically regard this as unethical or evil.

However, the moral debate continues to pop up from time to time and especially in connection with large scandals or crises. The largest one in recent memory was of course the global financial crisis of 2008, which can only be compared in size to the Great Depression of the 1930s. Part of the cause of the crisis seems to have been exactly a lack of ethics or social responsibility on the part of financial agents: for example, widespread speculation at very high risk, use of predatory lending practices, misleading or misguided financial "innovations," and excessive compensation programs. The crisis not only threatened the whole financial system with collapse, but also led to a broader economic recession that put many ordinary citizens' welfare at risk.

This chapter seeks to give an overview of the ethical issues and debates that pertain to money and finance, and thereby to help further establish the field of "financial ethics." Although a growing number of philosophers and critical thinkers have become interested in these issues, especially since the financial crisis, it seems fair to say that the field

is young and still searching for a common ground or intellectual meeting place. Part of the problem may be the lack of an established canon or a systematic overview of the field. This is exactly what this chapter seeks to remedy. However, it should be noted that constraints of time and space make it difficult to cover each and every issue and view. The most important limitation of the chapter is that it does not address the political dimension or the role of the state in the context. This is a major limitation since many of the debates touched upon have an obvious political dimension.

The chapter proceeds as follows. Section 17.2 gives some background by way of introducing ontological ideas about what money and finance is. Section 17.3 then plunges into the ethical issues by discussing some of the classic and sweeping criticisms to the effect that all, or at least most, financial activities are morally suspect. Section 17.4 assumes that the existence of financial markets can be acceptable and discusses some of the ethical issues involved in making them honest and fair for all parties. Finally, Section 17.5 discusses ideas to the effect that financial agents have social responsibilities that go beyond their role as market participants, such as an extended responsibility to promote social welfare. The chapter closes with concluding remarks on the importance of financial ethics.¹

17.2 WHAT IS MONEY AND FINANCE?

In order to adequately analyze the ethical issues raised by money and finance, we first need to understand their basic ontology. Money is so ever-present in modern life that we tend to take its existence and nature for granted. However, it is interesting how easily it can become elusive once we try to understand or define it more exactly. Do we really know what money is?

17.2.1 What Is Money?

The commodity theory of money. A classic and seemingly intuitive theory holds that money is just a special kind of commodity. More precisely, it is any commodity that fulfills three functions: of being a medium of exchange, a unit of account, and a store of value. Imagine a society that lacks money and in which people have to barter goods with each other. Barter only works when there is a *double coincidence of wants*: that is, when A wants what B has and B wants what A has. But since such coincidences are likely to be uncommon, a barter economy seems both cumbersome and inefficient (Smith 1776; Menger 1892). At some point, people will realize that they can trade more easily if they

¹ This chapter extends and elaborates on the author's contribution to the forthcoming entry on "Philosophy of Money and Finance" in the *Stanford Encyclopedia of Philosophy*, written with Boudewijn de Bruin, Lisa Herzog, and Martin O'Neill.

use some intermediate good—money. This intermediate good should ideally be easy to handle, store, and transport to function well (as medium of exchange). It should be easy to measure and divide to facilitate calculations (unit of account). And it should be difficult to destroy so that it lasts over time (store of value).

Monetary history may be viewed as a process of improvement on these functions of money (Ferguson 2008; Weatherford 1998). For example, some early societies used basic necessities that everyone valued as money, such as cattle or grain. Other societies settled on commodities that were easier to handle and to tally but with more indirect value, such as clam shells and precious metals. The archetypical money throughout history are gold or silver coins; therefore the commodity theory is sometimes called *metallism* (Knapp 1924; Schumpeter 1954). Coinage is an improvement on bullion in that both quantity and purity are guaranteed by some third party, typically the state. Finally, paper money can be viewed as a simplification of the trade in coins. For example, a bank note issued by the Bank of England in the 1700s was a promise to pay the bearer a certain pound weight of sterling silver.

The commodity theory of money was defended by many classical economists and can still be found in most economics textbooks (Mankiw 2009; Parkin 2011). This latter fact is curious since there has been serious and sustained critique of the theory. An obvious challenge is that it has difficulties in explaining inflation (the decreasing value of money over time), especially when the prices of metals are stable (Innes 1913). Another challenge is that the theory seems to be wrong about history. For example, recent anthropological studies question the idea that early societies went from a barter economy to money; instead, money seems to have arisen to keep track of pre-existing credit relationships (Graeber 2011; Martin 2013).

The credit theory of money. According to the main rival theory, then, coins and notes are merely tokens of something more abstract. One could say that money is a social construction rather than a physical commodity (Searle 1995). The abstract entity in question is a credit relationship: that is, a promise from someone to grant (or repay) a favor (product or service) to the holder of the token (Macleod 1889; Innes 1914). In order to function as money, two further features are crucial: that the promise is sufficiently credible, which means that the issuer is “creditworthy,” and that the credit is transferable, which means that also others will accept it as payment for trades. The latter is often a function of the former: a token will be transferable precisely when its issuer is regarded as creditworthy.

It is commonly thought that the most creditworthy issuer of money is the state. Here is then a different explanation of the predominance of coins and notes whose value is guaranteed by states. But note that the new theory also can explain so-called fiat money, money that is underwritten by the state but not redeemable in any commodity like gold or silver. Fiat money has been the dominant kind of money globally since 1971, when the United States terminated the convertibility of dollars to gold, which effectively made most national currencies free-floating. The view that only states can issue money is called *chartalism*, or the state theory of money (Knapp 1924). However, in order to properly understand the current monetary system, it is important to distinguish between

states' *issuing* versus *underwriting* money. Most credit money in modern economies is actually issued by commercial banks through their lending operations, and the role of the state is only to guarantee the convertibility of bank deposits into cash (Pettifor 2014).

Criticisms of the credit theory tend to focus exactly on the risk of *overexpansion* of money: that is, that states (and banks) can overuse their "printing presses" which may lead to unsustainable debt levels, financial instability, and economic crises. These are sometimes seen as arguments for a return to the gold standard (Rothbard 1983; Schlichter 2014). However, others argue that the realization that money is socially constructed is in fact the best starting point for developing a more sustainable and equitable monetary regime (Graeber 2010; Pettifor 2014).

17.2.2 What is Finance?

One may view "finance" more generally, the financial sector or system, as basically an extension of the monetary system. It is typically said that the financial sector has two main functions: (1) to maintain an effective payments system (to make sure that money is readily available and transferable in the economy), and (2) to facilitate an efficient use of the money (that is, to channel funds to those individuals or companies who can use it in the most productive ways). The latter and more ambitious function can be broken down further into three parts:

- First, to bring together those with excess money (savers, investors) and those without it (borrowers, enterprises). This is typically done through financial markets (such as stock or bond markets) or financial intermediation (the inner workings of banks).
- Second, to make it sufficiently attractive for the participants to buy and sell the money. This is typically done through the invention of financial products, or "assets," with suitable features (such as reduced risk, shorter maturities, or higher returns).
- Third, to adequately signal where money can be put to most productive use. This is typically done through facilitating information exchange and the pricing of products based on the expected profitability of underlying enterprises.

Financial assets. Modern finance thus consists of several other "asset types" besides money, which are of interest from an ontological viewpoint. Some central examples include credit arrangements (bank accounts, bonds), equity (shares or stocks), derivatives (futures, options, swaps, and so on) and funds (trusts). What is the defining characteristic of a financial asset?

The typical distinction here is between financial and "real" assets, such as buildings and machines (Fabozzi 2002). This does not mean that financial assets are unreal, but simply that they are less tangible or concrete. Just like money, they can be viewed as a social construction. A complication is that financial assets often derive from or at least

involve underlying “real” assets—compare, for example, between buying a house and investing in a housing company. In this sense, one may view the financial market as the “meta-level” of the economy, because it involves indirect trade or speculation on the success of other parts of the economy (that form the underlying basis for financial assets). This is indeed the reason for why financial markets can help to spread economic risks more thinly in society.

More distinctly, financial assets are defined as promises of future money payments (Mishkin 2016; Pilbeam 2010). If the credit theory of money is correct, they can then be regarded as meta-promises, or promises on promises. The level of abstraction can sometimes become enormous. For example, a “synthetic collateralized debt obligation” (a form of derivative common before the financial crisis) is a promise from person A (the seller) to person B (the buyer) that some persons C to I (speculators) will pay an amount of money depending on the losses incurred by person J (the holder of an underlying derivative), which typically depend on certain portions (so-called tranches) of the cash flow from persons K to Q (mortgage borrowers) originally promised to persons R to X (mortgage lenders) but then sold to person Y (the originator of the underlying derivative).

Intrinsic value. Perhaps the most important characteristic of financial assets is that their *price* is less tangible and can therefore vary enormously. Put simply, there are two main factors that determine the price of a financial asset (which are the same as in the credit theory of money above): (i) the credibility or prosperity of the underlying promise (the issuer’s future cash flow), and (ii) its transferability or popularity on the market (how many other investors that are interested in buying the asset). The process known as “price discovery” is when investors assess these factors based on the information available to them and then make bids to buy or sell the asset which in turn sets its price on the market (Mishkin 2016; Pilbeam 2010).

A philosophically interesting question is whether there is such a thing as an “intrinsic” or “true” price of financial assets. This is often assumed in discussions about financial crises. For example, a common definition of an “asset bubble” (such as a bubble in real estate bonds) is that the asset trades at a price that strongly exceeds its intrinsic value—which is dangerous because the bubble can burst and cause an economic shock (Kindleberger 1978; Minsky 1986; Reinhart and Rogoff 2009). But what then is the intrinsic value of an asset? A common answer is that this depends only on “fundamental” factors concerning the underlying “real” cash flow—in other words, on (i) and not (ii) above. However, someone still has to assess these factors to compute a price, which inevitably is a subjective process. As just noted, it is assumed that different investors have different valuations of financial assets (which is why they engage in trades on the market).

A further complication here is that (i) may actually be influenced by (ii)—that is, the fundamentals may be influenced by investors’ perceptions of them, a phenomenon known as “reflexivity” (Soros 1987, 2008). For example, a company whose shares are popular among investors will often find it easier to borrow more money and thereby to expand its cash flow, in turn making it even more popular among investors. Conversely, when the company’s profits start to fall it may lose popularity among investors, thereby

making its loans more expensive and its profits even lower. This phenomenon amplifies the risks posed by financial bubbles.

17.3 MONEY AS THE ROOT OF ALL EVIL?

Let us now turn to the broad range of compelling ethical issues related to money and finance. We start with the classic criticisms which suggest that all, or at least most, financial activities are morally suspect. We will here discuss three such criticisms, respectively directed at the love of money (the profit motive), usury (lending at interest), and speculation (gambling in finance).

17.3.1 The Love of Money

At the heart of many sweeping criticisms of money and finance lies an idea about motive. For example, the full Biblical quote is that “the love of money is the root of all [kinds of] evil” (1 Timothy 6:10). To have a “love of money” here means to desire or seek money for its own sake, or as an end in itself rather than as a means to something else. In modern times, we often speak of the profit motive of commercial agents. This signifies their over-arching goal of collecting revenues in excess of costs or, more simply, to “make money.” The profit motive has then been the subject of much moral criticism throughout history and continues to be controversial in popular morality.

But why is this so? There are three main variations of the criticism. A first variation says that there is something unnatural about the profit motive itself. For example Aristotle (330 BCE) argued that we should treat objects in a way that is befitting to their fundamental nature, and because money is not meant to be a good in itself but only a medium of exchange, he concluded that it is unnatural to desire money as an end in itself. A similar thought is picked up by Marx (1867), who argued that capitalism replaces the natural economic cycle of C–M–C (commodity exchanged for money exchanged for commodity) with M–C–M (money exchanged for commodity exchanged for money). Thus the endless accumulation of money becomes the sole goal of the capitalist, which Marx described as a form of “fetishism.” Indeed, moneylenders reduce the cycle even further to M–M (money exchanged for money), which is even more unnatural.

A second variation of the criticism concerns the broader character, or more precisely vice, that the profit motive is thought to exemplify (alternatively to stem from or lead to). To have a love for money is typically associated with selfishness and greed, that is, a desire to have as much as possible for oneself, or more than one really needs (McCarty 1988; Walsh and Lynch 2008). Another association is the loss of moral scruples so that one is ready to do anything for money—even to sell one’s soul to the devil, as the popular saying goes. Once again the financial industry is often held out as the worst

in this regard. A reason for this can be the extremely high levels of compensation in the industry, which far exceed those of other professions (Hendry 2013). Allegations of greed soared after the 2008 crisis, when financial executives continued to receive million-dollar bonuses while many ordinary workers lost their jobs (McCall 2010).

A third variation of the criticism says that the profit motive simply signals the absence of other and more appropriate motivations. For example, Kant (1785) argued that actions only have moral worth if they are performed for moral reasons, or, more specifically, for the sake of duty. Thus, it is not enough that we do what is right, but we must also do it *because* it is right. Another relevant Kantian principle is that we never should treat others merely as means for our own ends, but always also as ends in themselves. Both of these principles seem to contrast with the naked and instrumental profit motive of many commercial agents. Although this may not mean that they are immoral, they are at least not morally optimal (Bowie 1999; Maitland 2002).

Present-day capitalist societies obviously exhibit a more positive view toward commerce and the profit motive than this. There are two main lines of defensive argumentation in the literature. The most influential is Adam Smith's (1776) well-known argument about the positive side-effects in society of a self-interested pursuit of profits: although the baker and brewer only aim at their own respective good, Smith suggested, they are "led by an invisible hand"—economic competition—to at the same time promote the public good. This argument is typically viewed as a utilitarian vindication of the profit motive and commercial activity. That is, the point is that positive societal effects can morally outweigh the possible shortcomings in individual virtue (Flew 1976).

A second argument is more direct and holds that the profit motive can exemplify a positive virtue. For example, there is the well-known Protestant work ethic that emphasizes the positive nature of hard work, discipline, and frugality (Long 1972; Wesley 1960). The profit motive can, on this view, be associated with virtues such as ambition, industry, and discipline. According to Weber (1905), the Protestant work ethic played an important role in the development of capitalism as such.

17.3.2 Usury and Interest

If having a love of money seems morally suspect, then the practice of making money on money—for example, lending money at interest—could seem even worse. This is another sweeping criticism directed at finance that can be found from the traditional ethicists. Societies in both Ancient and Medieval times typically condemned or banned the practice of "usury," which originally simply meant the charging of interest on loans. As the practice started to become socially acceptable over time, usury came to mean the charging of excessive rates of interest. However, modern Islam still contains a general prohibition against interest, and many countries still have at least partial usury laws, most often setting an upper limit on acceptable rates.

What could be wrong with lending at interest? Some of the more obscure arguments concern (again) the nature of money: Aristotle (330 BCE) argued that there is something

unnatural with “money begetting money.” While money may be a useful means for facilitating commercial exchange, he thought that it has no productive use in itself and so receiving interest over and above the borrowed amount is unnatural and wrong. A somewhat related argument can be found in Aquinas (1485), who argued that money is a non-durable or consumable good: that is, a good that is consumed on use. Although a lender can legitimately demand repayment of an amount equivalent to the loan, then, it is illegitimate to demand payment for the use of the borrowed amount and so adding interest is unnatural and wrong. (One can only wonder what these authors would have thought about present-day financial derivatives, which are bets to make money on other bets to make money.)

Some more promising arguments concern justice and inequality. For example, Plato (380 BCE) expressed worries that allowing interest may lead to societal instability. It may be noted that the Biblical condemnations of usury most straightforwardly prohibit interest-taking from the poor, which can be given several interpretations. One idea is that we simply owe a duty of charity to the poor and charging interest is incompatible with this duty; we should instead give them what they need. Another idea is that the problem lies in the outcome of interest payments: Loans are typically extended by someone who is richer (someone with capital) to someone who is poorer (someone without it) and so asking for additional interest may increase the inequitable distribution of wealth in society (Sandberg 2012; Visser and MacIntosh 1998). A third idea, which is prominent in the Protestant tradition, is that lending often involves opportunism or exploitation in the sense of offering bad deals to poor people that effectively have no other options available (Graafland 2010).

The Islamic condemnation of interest, or *riba*, has roots in both of these strands of thought. But there is also a third line of argument which holds that interest essentially is unearned or undeserved income. Because the lender neither partakes in the actual productive use of the money lent, nor exposes himself or herself to commercial risk in the same way as the borrower does, the lender cannot legitimately share in the gains produced by the loan (Ayub 2007; Birnie 1952; Thomas 2006). Based on this argument, contemporary Islamic banks insist that lenders and borrowers must form a business partnership in order for fees on loans to be morally legitimate (Ayub 2007; Warde 2010). It is interesting to see how economists (also in the West) have struggled to find good retorts to this argument. Some economists stress that lending also involves risk (for example, that the borrower defaults and is unable to repay); others stress the so-called opportunity costs of lending (the money could have been used more profitably elsewhere); and yet again others stress the simple time-preference of individuals (that we value consumption in the present more than the future, and therefore the lender deserves compensation for postponing consumption).

The gradual abandonment of the Medieval usury laws in the West is typically attributed to a growing acknowledgement of the great potential for economic growth unleashed by easy access to loan capital. One could perhaps say that history itself disproved Aristotle: money (or at least easy access to it) indeed proved to have a productive use. In a short text from 1787, Bentham famously poked fun at many of the classical anti-usury arguments

and defended the practice of charging interest from a utilitarian standpoint. However, this does not mean that worries about and allegations of usury have disappeared entirely in society. As noted above, usury today means charging interest rates that seem excessive or exorbitant. For example, many people are outraged by the rates charged on modern payday loans, or the way in which rich countries exact interest on their loans from poor countries (Baradaran 2015; Herzog 2017). These intuitions have clear affinities with the justice-based arguments outlined above.

17.3.3 Speculation and Gambling

The most contemporary sweeping criticism of money and finance concerns the supposed moral defects of speculation. This criticism tends to be directed towards financial activities that go beyond mere lending. Critics of the capitalist system often liken the stock market to a casino and investors to gamblers or punters (Sinn 2010; Strange 1986). More moderate critics insist on a strict distinction between investors or shareholders, on the one hand, and speculators or gamblers, on the other (Bogle 2012; Sorell and Hendry 1994). In any case, the underlying assumption is that the similarities between modern financial activities and gambling are morally troublesome. Unfortunately, this line of criticism is also quite obscure and can be given several interpretations.

On some interpretations, the criticism is similar to concerns raised above. For example, some argue that speculators are driven by a naked and relentless profit motive whereas investors have a genuine concern for the underlying business enterprise. The immorality of speculation would then consist in the inappropriateness of the profit motive, or the vices of selfishness and greed (Hendry 2013). Others liken investment with gambling because they view the former as “parasitic,” to be without productive use and solely dependent on luck (Borna and Lowry 1987; Ryan 1902). This argument is similar to the complaint about undeserved income discussed above. It should come as no surprise that modern Islam is also critical of financial speculation (Ayub 2007; Warde 2010).

A more distinct interpretation holds that speculation typically includes very high or excessive levels of risk-taking (Borna and Lowry 1987). This is morally problematic when the risks not only affect the gambler him- or herself but also communities or even entire societies. For example, a root cause of the financial crisis of 2008 was widespread speculation on very risky derivatives such as synthetic collateralized debt obligations (described above). When the value of such derivatives fell dramatically, the financial system as a whole came to the brink of collapse. (We will return to this criticism below.)

A related interpretation concerns the supposed short-sightedness of speculation. It is often argued that financial agents and markets are “myopic” in the sense that they only care about profits in the very near term, often the next quarter (Dallas 2012). The myopia can stem from agents’ supposedly greedy nature or simply the time-preference of individuals noted above. Alternatively it is connected with modern disclosure requirements that force companies to publish quarterly earnings reports (on which more below).

In any case, the myopia of finance is typically blamed for negative effects such as market volatility, the continuous occurrence of manias and crashes, inadequate investment in social welfare, and the general short-sightedness of the economy (Lacke 1996).

Defenders of speculation argue that it can serve a number of positive ends. To the extent that all financial activities are speculative in some sense, of course, the ends coincide with the function of finance more generally: to channel funds to the individuals or companies who can use them in the most productive ways. But even speculation in the narrower sense—of high-risk, short-term bets—can have a positive role to play. It can be used to “hedge” or off-set the risks of more long-term investments, and it contributes to “market liquidity” (roughly, the amount of counterparties to trade with at any given point of time), which is important for an efficient pricing mechanism (Angel and McCabe 2009; Koslowski 2009).

17.4 FAIRNESS IN FINANCIAL MARKETS

Let us now assume that the existence of financial markets can be ethically acceptable, at least in principle, so that we can turn to discuss some of the issues involved in making them fair and just for all parties. We will focus on three such issues here, respectively concerned with deception and fraud (honesty), conflicts of interest (care for customers), and insider trading (fair play).

17.4.1 Deception and Fraud

Some of the most publicized ethical scandals in finance are examples of deception or fraud. For example, Enron (a huge US corporation) went bankrupt after it was discovered that its top managers had “cooked the books”—that is, engaged in fraudulent accounting practices. They had basically kept huge debts off the company’s balance sheet in an effort to make it look more profitable and thereby boost the share price. Other scandals in the industry have involved deceptive marketing practices, hidden fees or costs, undisclosed or misrepresented financial risks, and outright Ponzi schemes.

While these examples seem obvious, on further examination it is not easy to give an exact definition of financial deception or fraud—and, of course, proving it in practice is even harder. The most straightforward case seems to be deliberately misrepresenting or lying about financial facts. However, this assumes that there is such a thing as a financial fact, that is, a correct way of understanding or representing a financial value or transaction. In light of the socially constructed nature of money and finance, this is far from clear. Less straightforward cases include simply concealing or omitting financial information, or refraining from obtaining the information in the first place.

A philosophical conception of fraud, inspired by Kant, holds that it consists of denying the weaker party to a financial transaction (such as a consumer or investor) information

that is necessary to make a rational (or autonomous) decision (Boatright 2014; Duska and Clarke 2002). Many countries require that the seller of a financial product (such as a company emitting shares) must disclose all information that is “material” to the product, which may be understood along similar lines. It is an interesting question whether this suggestion, especially the conception of rationality involved, includes or rules out a consideration of the ethical nature of the product, such as the ethical nature of the company’s operations (Lydenberg 2014). Furthermore, there may be information that is legitimately excluded by other considerations, such as the privacy of individuals or companies commonly protected by “bank secrecy” laws.

But is access to adequate information enough? A complication here is that the weaker party, especially ordinary consumers, may have trouble in understanding or processing the information enough to identify cases of fraud. This is a structural problem in finance that has no easy fix. As we have seen, financial products are often abstract, complex, and difficult to price. The implication for theory is that full autonomy of agents may not only require access to adequate information, but also access to sufficient processing ability and resources to analyze the information (Boatright 2014). Some therefore argue that the weaker party must be given further help and protection by the state, for example through consumer protection laws or agencies (de Bruin 2015; Shiller 2012).

17.4.2 Avoiding Conflicts of Interest

Due to the problems just noted, the majority of ordinary consumers refrain from engaging in financial markets on their own and instead rely on the services of financial intermediaries, such as banks, investment funds, and insurance companies. For example, instead of investing directly in shares themselves, they rely on professional fund managers to have a better understanding of what shares to buy and sell. But this leads to closely related ethical problems that are due to the conflicts of interest inherent in financial intermediation. Simply put, the managers or employees of intermediaries have ample opportunity, as well as incentive (when there is a culture of greed), to misuse their customers’ money and trust.

Although it once again is difficult to give an exact definition, the literature is full of examples of such misuse, including so-called churning (trading excessively to generate high fees), stuffing (selling the bank’s undesired assets to a client), front-running (buying an asset for the bank first and then reselling it to the client at a higher price), and tailgating (mimicking a client’s trade to piggyback on his or her information) (Dilworth 1994; Heacock et al. 1987). Interestingly, some argue that the whole industry of actively managed investment funds may be seen as a form of fraud. According to economic theory, namely, it is impossible to beat the average returns of the market for any given level of financial risk, at least in the long term. Therefore, funds who claim that they can do this for a fee are basically cheating their clients (Hendry 2013).

A legal doctrine that aims to protect clients here is so-called fiduciary duty, which imposes obligations on fiduciaries (those entrusted with others’ money) to act in the

sole interest of beneficiaries (those who own the money). The interests referred to are typically taken to be financial interests, so the obligation of the fiduciary is basically to maximize investment returns. But some argue that there are cases in which beneficiaries' broader interests should take precedence, such as when investing in fossil fuels may give high financial returns but pose serious risks to people's future lives (Lydenberg 2014; Sandberg 2013, 2016). In any case, it is often thought that fiduciary duty goes beyond the ideal of a free market to instead give stronger protection to the weaker party of a fragile relationship, in this case the passive consumer.

As an alternative or compliment to fiduciary duty, some argue for the adoption of a code of ethics or professional conduct by financial professionals. A code of ethics would be less arduous in legal terms and is therefore more attractive to free market proponents (Koslowski 2009). It can also cover fragile relationships beyond that of fiduciary-beneficiary, including those of bank-depositor, advisor-client, and so forth. Just as doctors and lawyers have a professional code, finance professionals could have one that stresses values such as honesty, due care, and accuracy (de Bruin 2016; Graafland and Ven 2011). But according to critics, the financial industry is simply too divided (into separate roles and competencies) to have a uniform code of ethics (Ragatz and Duska 2010). It is also unclear whether finance can be regarded as a profession in the traditional sense, which typically requires a body of specialized knowledge, high degrees of organization and self-regulation, and a commitment to public service (Boatright 2014; Herzog 2017).

17.4.3 Insider Trading

Probably the most well-known ethical problem concerning fairness in finance—and also the one that philosophers disagree the most on—is so-called insider trading. Put simply, it is when an agent uses his or her position within, or privileged information about, a company to buy or sell its shares at favorable times and prices. For example, a CEO may buy shares in her company just before it announces a major increase in earnings that will boost the share price. It seems that no deception or fraud is involved in this case because there is no spreading of false information; nor is there any breach of duty (we may assume) towards the counterparty in the trade. However, a similarity with the previous examples is that the agent is exploiting an asymmetry of information.

Just like before, it is difficult to give an exact definition of insider trading. People in different positions in a company obviously have access to different information with varying relevance to financial markets. However, most commentators agree that it is the information that counts and, thus, the insider need not actually be inside the company—instead, it can be family, friends, or other tippees (Irvine 1987a; Moore 1990). Indeed, some argue that the information need not come from inside the company either: for example, stock analysts and journalists can be regarded as insiders if they trade on information that they have gathered themselves but not yet made available to a wider audience. It is also debatable whether an actual trade has to take place or whether insider

trading can consist in an omission to trade based on inside information, or also in enabling others to trade or not trade in this way (Koslowski 2009).

Several philosophical perspectives have been used to explain what (if anything) is wrong with insider trading. One perspective invokes the concept of fair play. Even in a situation with fully autonomous traders, the argument goes, market transactions are not fair if one party has access to information that the other has not. This can be likened to playing cards with a marked deck, where only one of the players knows about the markings. Fair play requires a “level playing field” in which no participant starts from an unfairly advantaged position (Werhane 1989, 1991). However, critics argue that this perspective imposes excessive demands of informational equality. There are many asymmetries of information in the market that are seemingly unproblematic; for example, antiquaries usually know more about antiques than their customers (Lawson 1988; Machan 1996). So is it the inaccessibility of inside information that is problematic? But one could argue that, in principle, outsiders have the possibility to become insiders and thus to obtain the exact same information (Lawson 1988; Moore 1990).

Another perspective views insider trading as a breach of duty, not towards the counterparty in the trade but towards the source of the information. For example, US legislation treats inside information as the property of the underlying company and, thus, insider trading is essentially a form of theft of corporate property, often called the misappropriation theory (Lawson 1988). A related suggestion is that it can be seen as a violation of the fiduciary duty that insiders have towards the company they work for (Moore 1990). However, critics argue that the misappropriation theory misrepresents the relationship between companies and insiders. On the one hand, there are many normal business situations in which insiders are permitted or even expected to spread inside information to outside sources (Boatright 2014). On the other hand, if the information is the property of the company, why do we not allow it to be “sold” to insiders as a form of remuneration? (On this, see Engelen and Liedekerke 2010; Manne 1966.)

A third perspective deals with the effects, both direct and indirect, of allowing insider trading on the market. Interestingly, many argue that the direct effects of such a policy are likely to be positive. As noted above, one of the main purposes of financial markets is to form (or “discover”) prices that reflect all available information about a company, which in turn directs capital to its most productive use. Because insider trading contributes important information, it is likely to improve on this process of price discovery (Manne 1966). Indeed, the same reasoning suggests that insider trading helps the counterparty in the trade to get a better price, so it is a victimless crime (Engelen and Liedekerke 2010). However, others express concern over the indirect effects that are likely to be more negative. For example, allowing insider trading may erode the moral standards of market participants by favoring opportunism over fair play (Werhane 1989). Moreover, many people may be dissuaded from even participating in the market if they feel that it is “rigged” to their disadvantage (Strudler 2009).

17.5 THE SOCIAL RESPONSIBILITY OF FINANCE

Let us finally take a broader societal view on finance and discuss claims that financial agents have social responsibilities that go beyond their basic role as market participants. We will discuss three such ideas here, respectively focusing on systemic risk (responsibility to avoid societal harm), microfinance (responsibility towards the poor or unbanked), and socially responsible investment (responsibility to address a range of societal challenges).

17.5.1 Systemic Risk and Financial Crises

A root cause of the financial crisis of 2008 was the very high levels of risk-taking of many banks and other financial agents, such as bets on bonds and derivatives tied to the US real estate market. As noted above, when these bets failed dramatically, the financial system as a whole came to the brink of collapse. Many banks lost so much money that their normal lending operations were hampered, which in turn had negative effects on the real economy and millions of “ordinary” citizens around the world lost their jobs. In the end, many governments had to step in to bail out the banks and thereby sacrifice other parts of public spending, causing further stress to the average citizen. This is a prime example of how certain financial activities, when run amok, can have devastating effects on third parties and society in general.

It is important to note that the risk-taking may have been prudent for each bank in isolation, only taking into account the probability of losses for that particular bank. This is the only focus of standard definitions of financial risk (Thamotheram and Ward 2014). However, it is also clear that the aggregate or collective level of risk was unsound. In order to capture this, much subsequent debate has focused on so-called systemic risk, that is, the risk of failures across several agents which impairs the functioning of the financial system as such, including basic services such as lending (Brunnermeier and Oehmke 2013; Smaga 2014). An interesting distinction in this context is that between prudence (on the individual level) and “macroprudence” (on the collective).

The concept of systemic risk gives rise to several prominent ethical issues. For example, to what extent do financial agents have a moral duty to limit their contributions to systemic risk? It could be argued that financial transactions always carry risk and that this is “part of the game.” But the important point about systemic risk is that financial crises not only affect the transacting parties themselves but, as just noted, have negative effects on third parties. This constitutes a *prima facie* case for a duty of precaution on the part of financial agents, based on the social responsibility to avoid causing unnecessary harm (James 2017; Linarelli 2017). This *prima facie* case may be supplemented with mitigating conditions, such as that the strength of the duty depends on its cost to the agent

and her access to (information about) adequate alternatives (Linarelli 2017). In cases when precaution is impossible, one could add a related duty of rectification or compensation to the victims of the harm (James 2017).

Two factors determine how much an agent's activity contributes to systemic risk (Brunnermeier and Oehmke 2013; Smaga 2014). The first is financial risk in the traditional sense: that is, the probability and size of the potential losses for that particular agent. High financial risk translates into a high probability of serious losses or even default in the event of an unexpected financial shock (such as a dramatic fall in real estate prices). A duty of precaution may here be taken to imply, for example, stricter requirements on capital and liquidity reserves (roughly, the money that the agents must keep in their coffers for emergency situations). The second factor is the agent's place in the broader financial system, which typically is measured by its interconnectedness with other agents: that is, to what extent other agents financially depend on its business. High interconnectedness means that the agent's possible losses or default could have cascading implications for many other agents as well, a process known as contagion. Interestingly, the reasoning here seems to imply that the duty of precaution is stronger for financial agents that are "systemically important" or, as the saying often goes, "too-big-to-fail" institutions.

As an alternative to the reasoning above, one may argue that the duty of precaution is more properly located on the collective or political level (James 2012, 2017). Systemic risk is, after all, caused by the collective behavior of market participants and no individual agent will be able to avert a catastrophe on his or her own. As noted at the outset, many of the issues here have an obvious political dimension that go beyond this chapter.

17.5.2 Microfinance

Even in the absence of financial turmoil, some people with very low income or wealth have limited access to basic financial services such as loans. The problem here seems to be that commercial banks have little to gain from offering them such services; there is an elevated risk of loan losses (because the poor lack collateral) and it is costly to administer a large amount of very small loans (Armendáriz and Morduch 2010). Moreover, one cannot rule out that some bank officers discriminate against underprivileged groups. An initiative that seeks to remedy these problems is the so-called microfinance (or microcredit) movement. Microfinance is exactly the extension of financial services, such as lending and saving, to poor or low-income people who are otherwise "unbanked." This often takes place in some of the poorest countries of the world, such as Bangladesh and India.

The justifications offered for microfinance are similar to those for development aid. For example, a popular justification holds that affluent people have a duty of assistance towards those less fortunate, and microfinance is thought to be a particularly efficient way to alleviate poverty (Yunus 1998, 2007). But is this correct? Judging from the growing number of empirical "impact studies," it seems more correct to say that microfinance

sometimes is helpful, but at other times can be either ineffective or have negative side effects (Hudon and Sandberg 2013; Roodman 2012). Another justification holds that there is a basic human right to subsistence, and this includes a right to important economic means such as savings and credit. The relevant right is associated with an obligation on the part of financial institutions to provide affordable services, as well as an obligation on states to assist in this matter (Hudon 2009; Meyer 2018). But would this work? Critics argue that the framework of human rights is too rigid and absolute to be a good fit for financial services, which typically come with both benefits and challenges (Gershman and Morduch 2015; Sorell 2015).

Of course, microfinance is ultimately different from development aid in that it involves commercial banking relations. This feature invites the familiar political debate of state- versus market-based society and development. Proponents of microfinance tend to argue that traditional state-led development projects have been too rigid and corrupt, whereas market-based initiatives are more flexible and help people to help themselves (Armendáriz and Morduch 2010; Yunus 2007). According to critics, however, it is the other way around: markets will tend to breed greed and inequality, whereas real development is created by large-scale investments in education and infrastructure (Bateman 2010; Weber 2004).

In recent years, the microfinance industry has seen several ethical scandals that seemingly testify to the risk of market excesses. For example, reports have indicated that interest rates on microloans average 20–30 percent per annum, and can sometimes be in excess of 100 percent, which is much higher than the rates for non-poor borrowers. This raises questions about usury, which as we have seen have a long philosophical history, as the high rates are often taken to be exorbitant or exploitative (Hudon and Ashta 2013; Rosenberg et al. 2009). However, some suggest a defense of “second best,” or last resort, when other sources of aid or cheaper credit are generally unavailable (Sandberg 2012). Microfinance institutions have also been accused of using coercive lending techniques and forceful loan recovery practices (Dichter and Harper 2007; Priyadarshiee and Ghalib 2012). This raises questions about the ethical justifiability of commercial activity directed at the desperately poor: even when there is a lack of physical coercion, very poor customers may have no viable alternative than to accept deals that are both unfair and exploitative (Arnold and Valentin 2013; Hudon and Sandberg 2013).

17.5.3 Socially Responsible Investment

Socially responsible investment refers to the emerging practice of financial agents to give weight to putatively ethical, social, or environmental considerations in investment decisions, including decisions about what bonds or stocks to buy or sell, and how to engage with the companies in one’s portfolio. This is sometimes part of a mainstream and strictly profit-driven investment philosophy, based on the assumption that companies with superior social performance also will have superior financial performance (Richardson and Cragg 2010). But more commonly, it is perceived as an alternative to

mainstream investment. The background argument here is roughly that the pricing mechanism of both real and financial markets seems unable to promote sufficient levels of social and environmental responsibility in firms. For instance, even though there is widespread social agreement on the evils of sweatshop labor and environmental degradation, mainstream investors are still financing such enterprises. Therefore, there is a need for a new kind of investor with a stronger sense of social responsibility (Sandberg 2008; Cowton and Sandberg 2012).

The simplest and most common practice among these alternative investors is to avoid investments in companies that are perceived to be ethically problematic. This is typically justified from a deontological idea to the effect that it is wrong to invest in someone else's wrongness; that is, that the ethically problematic nature of the underlying corporate activity spills over to, or "taints," also those who invest in it (Irvine 1987b; Langtry 2002; Larmer 1997). There are at least three interpretations of the basis for this moral taint. One is that the salient wrongness consists in profiting from wrongdoing, or perhaps in benefitting from other people's suffering. Another is that the wrongness consists simply in harming others, or at least in facilitating harm. Finally, a third interpretation holds that the salient relationship is more of a symbolic nature, namely that of "morally supporting" or indicating one's acceptance of wrongful activities.

The deontological perspective above has been criticized for being too black-and-white. On the one hand, it seems difficult to find a single investment opportunity that is completely "pure" because modern corporations have very diverse operations, and also tend to be connected to each other through either operational or financial ties (Kolars 2001). On the other hand, the relationship between the investor and the investee is not as direct as one may think. For starters, not all companies pay dividends and it is not uncommon that investments lead to losses rather than profits for the investor. Furthermore, to the extent that investors buy and sell shares on the stock market, they are not engaging with the underlying companies but rather with other investors. The only way in which such transactions could benefit the companies would be through movements in the share price (which determines the so-called cost of capital), but it is extremely unlikely that a group of ethical investors can affect this. After all, the *raison d'être* of stock exchanges is exactly to create markets that are sufficiently liquid to maintain stable prices (Haigh and Hazelton 2004; Hudson 2005).

A rival perspective on socially responsible investment is the consequentialist idea that investors' duty towards society consists in using their financial powers (as well as formal powers as shareholders of portfolio companies) to promote positive societal goods, such as social justice and environmental sustainability. This perspective is typically taken to prefer more progressive investment practices, such as seeking out environmentally friendly technology firms or pushing management to adopt more ambitious social policies (Mackenzie 1997; Sandberg 2008). Of course, the flip side of such practices, which may explain why they are less common in the market, is that they invite greater financial risks. For example, investing in new technology is always very risky and may therefore be imprudent from a financial viewpoint, even when the potential

societal rewards are enormous (Sandberg 2011). It remains an open question whether socially responsible investment will grow enough in size to make financial markets a force for positive change in society.

17.6 CONCLUDING REMARKS

The present chapter has sought to give an overview of the many ethical issues and debates that pertain to money and finance. It should hopefully have become clear how money and finance are quite special phenomena and therefore also raise special ethical issues. Let me briefly comment on these specificities here. One special characteristic is the abstractness. Money is not just another commodity, but something more ethereal and universal that we use as a medium of exchange, and finance is an even more abstract layer that allows us to exchange and trade in money itself. This has led some philosophers to argue that money-loving motives are unnatural or unethical, and that all financial activities are morally suspect. Although such a reaction may be excessive, we have seen that financial activities indeed can be obscure and easily manipulated. This suggests that financial agents have a special responsibility to make sure that their marketing, managing, and trading of financial products live up to moral standards. Such a responsibility is especially important in relation to weaker clients, such as middle- and low-income consumers, who we have seen are doubly vulnerable to being exploited by the financial system: they have the least money and therefore the least sway in the system, and they often lack sufficient knowledge or interest to understand its abstract and complex workings.

A second special characteristic of money and finance is their great importance to society. Money is the glue that holds the economy together, one could say, and finance is the lubricant that keeps the wheels of the economy in motion. This suggests that financial agents have a special responsibility to understand the role and effects of their activities on third parties and society in general. We have seen that excessive complexity coupled with extreme risk-taking (without sufficient precaution) can lead to financial crises that put people's jobs and savings in jeopardy. Moreover, financial agents may have a social obligation to direct their lending and investing to causes or individuals that otherwise receive insufficient attention by other parts of society. There are additional questions here about the proper role of politics and regulation that we have been unable to address in the chapter.

The considerations above both speak to the great importance of the field of "financial ethics." Although a growing number of philosophers and critical thinkers have become interested in these issues, as noted at the outset of the chapter, it seems fair to say that the field is young and still searching for a common ground or intellectual meeting place. It is hoped that this chapter, through at least providing a systematic overview of the relevant issues, will spark further interest and debate in the field in the future.

ACKNOWLEDGMENTS

The author gratefully acknowledges financial support from the University of Gothenburg, the Knut and Alice Wallenberg Foundation, the Marcus and Amalia Wallenberg Foundation, and the Swedish Foundation for Strategic Environmental Research (Mistra).

REFERENCES

- Angel, James J., and Douglas M. McCabe. 2009. "The Ethics of Speculation." *Journal of Business Ethics* 90: 277–86.
- Aquinas, Thomas. 1485. *Summa Theologica*. Basel: Michael Wenssler.
- Aristotle. 330 BCE. *Politics*. Ernest Baker (trans.). Oxford: Oxford University Press (1995 edition).
- Arnold, Denis G., and Andres Valentin. 2013. "Corporate Social Responsibility at the Base of the Pyramid." *Journal of Business Research* 66: 1904–14.
- Armendáriz, Beatriz, and Jonathan Morduch. 2010. *The Economics of Microfinance*. 2nd edn. Cambridge, MA: MIT Press.
- Ayub, Muhammad. 2007. *Understanding Islamic Finance*. Chichester, UK: John Wiley & Sons.
- Baradaran, Mehrsa. 2015. *How the Other Half Banks: Exclusion, Exploitation, and the Threat to Democracy*. Cambridge, MA: Harvard University Press.
- Bateman, Milford. 2010. *Why Doesn't Microfinance Work?* London: Zed Books.
- Bentham, Jeremy. 1787. *Defence of Usury*. London: Payne and Foss.
- Birnie, Arthur. 1952. *The History and Ethics of Interest*. London: William Hodge & Company.
- Boatright, John R. 2014. *Ethics in Finance*. 3rd edn. Chichester, UK: John Wiley & Sons.
- Bogle, John C. 2012. *The Clash of Cultures: Investment vs. Speculation*. Hoboken, NJ: John Wiley & Sons.
- Borna, Shaheen, and James Lowry. 1987. "Gambling and Speculation." *Journal of Business Ethics* 6: 219–24.
- Bowie, Norman. 1999. *Business Ethics: A Kantian Perspective*. Oxford: Blackwell.
- Brunnermeier, Marcus K., and Martin Oehmke. 2013. "Bubbles, Financial Crises, and Systemic Risk." In George M. Constantinides, Milton Harris, and Rene M. Stulz (eds.), *Handbook of the Economics of Finance*, vol. 2B (Amsterdam: North Holland), pp. 1221–88.
- Cowton, Christopher J., and Joakim Sandberg. 2012. "Socially Responsible Investment." In Ruth Chadwick (ed.), *Encyclopedia of Applied Ethics*, 2nd ed., vol. 4 (San Diego, CA: Academic Press), pp. 142–51.
- Dallas, Lynne. 2012. "Short-Termism, the Financial Crisis, and Corporate Governance." *Journal of Corporation Law* 37: 265–364.
- de Bruin, Boudewijn. 2015. *Ethics and the Global Financial Crisis*. Cambridge: Cambridge University Press.
- de Bruin, Boudewijn. 2016. "Pledging Integrity: Oaths as Forms of Business Ethics Management." *Journal of Business Ethics* 136: 23–42.
- Dichter, Thomas, and Malcolm Harper (eds.). 2007. *What's Wrong with Microfinance?* Rugby, UK: Practical Action Publishing.
- Dilworth, John B. 1994. "The Ethical Importance of Conflicts of Interest: Accounting and Finance Examples." *Business & Professional Ethics Journal* 13: 25–40.
- Duska, Ronald F., and James J. Clarke. 2002. "Ethical Issues in Financial Services." In Norman E. Bowie (ed.), *The Blackwell Guide to Business Ethics* (Oxford: Blackwell), pp. 206–24.

- Engelen, Peter-Jan, and Luc Van Liedekerke. 2010. "Insider Trading." In John R. Boatright (ed.), *Finance Ethics: Critical Issues in Theory and Practice* (Hoboken, NJ: John Wiley & Sons), pp. 199–221.
- Fabozzi, Frank J. 2002. *The Handbook of Financial Instruments*. Hoboken, NJ: John Wiley & Sons.
- Ferguson, Niall. 2008. *The Ascent of Money: A Financial History of the World*. London: Penguin.
- Flew, Antony. 1976. "The Profit Motive." *Ethics* 86: 312–22.
- Gershman, John, and Jonathan Morduch. 2015. "Credit Is Not a Right." In Tom Sorell and Luis Cabrera (eds.), *Microfinance, Rights and Global Justice* (Cambridge: Cambridge University Press), pp. 14–26.
- Graafland, Johan J. 2010. "Calvin's Restrictions on Interest: Guidelines for the Credit Crisis." *Journal of Business Ethics* 96: 233–48.
- Graafland, Johan J., and Bert W. van de Ven. 2011. "The Credit Crisis and the Moral Responsibility of Professionals in Finance." *Journal of Business Ethics* 103: 605–19.
- Graeber, David. 2011. *Debt: The First 5000 Years*. New York: Melville House.
- Haigh, Matthew, and James Hazelton. 2004. "Financial Markets: A Tool for Social Responsibility?" *Journal of Business Ethics* 52: 59–71.
- Heacock, Marian V, Kendall P. Hill, and Seth C. Anderson. 1987. "Churning: An Ethical Issue in Finance." *Business and Professional Ethics Journal* 6: 3–17.
- Hendry, John. 2013. *Ethics and Finance: An Introduction*. Cambridge: Cambridge University Press.
- Herzog, Lisa. 2017. *Just Financial Markets? Finance in a Just Society*. Oxford: Oxford University Press.
- Hudon, Marek. 2009. "Should Access to Credit Be a Right?" *Journal of Business Ethics* 84: 17–28.
- Hudon, Marek, and Arvind Ashta. 2013. "Fairness and Microcredit Interest Rates: From Rawlsian Principles of Justice to the Distribution of the Bargaining Range." *Business Ethics: A European Review* 22: 277–91.
- Hudon, Marek, and Joakim Sandberg. 2013. "The Ethical Crisis in Microfinance: Issues, Findings, and Implications." *Business Ethics Quarterly* 23: 561–89.
- Hudson, Richard. 2005. "Ethical Investing: Ethical Investors and Managers." *Business Ethics Quarterly* 15: 641–57.
- Innes, Mitchell. 1913. "What is Money?" *Banking Law Journal* 30: 377–408.
- Innes, Mitchell. 1914. "The Credit Theory of Money." *Banking Law Journal* 31: 151–68.
- Irvine, William B. 1987a. "Insider Trading: An Ethical Appraisal." *Business and Professional Ethics Journal* 6 (4): 3–33.
- Irvine, William B. 1987b. "The Ethics of Investing." *Journal of Business Ethics* 6: 233–42.
- James, Aaron. 2012. *Fairness in Practice: A Social Contract for a Global Economy*. Oxford: Oxford University Press.
- James, Aaron. 2017. "The Distinctive Significance of Systemic Risk." *Ratio Juris* 30: 239–58.
- Kant, Immanuel. 1785. *Grundlegung zur Metaphysik der Sitten*. Riga: J. F. Hartknoch.
- Kindleberger, Charles P. 1978. *Manias, Panics, and Crashes: A History of Financial Crises*. London: Macmillan.
- Knapp, Georg Friedrich. 1924. *The State Theory of Money*. London: Macmillan.
- Kolers, Avery. 2001. "Ethical Investing: The Permissibility of Participation." *Journal of Political Philosophy* 9: 435–52.
- Koslowski, Peter. 2009. *Ethik der Banken: Folgerungen aus der Finanzkrise*. Munich: Wilhelm Fink Verlag.

- Lacke, Jay C. 1996. "The Ethics of Financial Derivatives in the History of Economic Thought." In W. Michael Hoffman, Judith Brown Kamm, Robert E. Frederick, and Edward S. Petry (eds.), *The Ethics of Accounting and Finance* (Westport, CT: Quorum Books), pp. 165–79.
- Langtry, Bruce. 2002. "The Ethics of Shareholding." *Journal of Business Ethics* 37: 175–85.
- Larmer, Robert. 1997. "The Ethics of Investing: A Reply to William Irvine." *Journal of Business Ethics* 16: 397–400.
- Lawson, Gary. 1988. "The Ethics of Insider Trading." *Harvard Journal of Law & Public Policy* 11: 727–83.
- Linarelli, John. 2017. "Luck, Justice and Systemic Financial Risk," *Journal of Applied Philosophy* 34: 331–52.
- Long, John D. 1972. "The Protestant Ethic Reexamined." *Business Horizons* 15: 75–82.
- Lydenberg, Steven. 2014. "Reason, Rationality and Fiduciary Duty." In James P. Hawley, Andreas G.F. Hoepner, Keith L. Johnson, Joakim Sandberg, and Edward J. Waitzer (eds.), *Cambridge Handbook of Institutional Investment and Fiduciary Duty* (Cambridge: Cambridge University Press), pp. 287–99.
- Machan, Tibor R. 1996. "What is Morally Right with Insider Trading." *Public Affairs Quarterly* 10: 135–42.
- Mackenzie, Craig. 1997. *Ethical Investment and the Challenge of Corporate Reform*. Bath: University of Bath.
- Macleod, Henry Dunning. 1889. *The Theory of Credit*. London: Longmans, Green, and Co.
- Maitland, Ian. 2002. "The Human Face of Self-Interest." *Journal of Business Ethics* 38: 3–17.
- Mankiw, Gregory. 2009. *Macroeconomics*. 7th edn. New York: Worth Publishers.
- Manne, Henry G. 1966. *Insider Trading and the Stock Market*. New York: Free Press.
- Martin, Felix. 2013. *Money: The Unauthorised Biography*. London: Bodley Head.
- Marx, Karl. 1867. *Das Kapital: Kritik der Politischen Oekonomie*. Hamburg: Verlag von Otto Meissner.
- McCall, John J. 2010. "Executive Compensation." In John R. Boatright (ed.), *Finance Ethics: Critical Issues in Theory and Practice* (Hoboken, NJ: John Wiley & Sons), pp. 547–64.
- McCarty, Richard. 1988. "Business and Benevolence." *Business & Professional Ethics Journal* 7 (2): 63–83.
- Menger, Carl. 1892. "On the Origins of Money." *Economic Journal* 2: 239–55.
- Meyer, Marco. 2018. "The Right to Credit." *Journal of Political Philosophy* 26: 304–26.
- Mishkin, Frederic. 2016. *The Economics of Money, Banking, and Financial Markets*. 11th edn. New York: Pearson Education.
- Minsky, Hyman. 1986. *Stabilizing an Unstable Economy*. New Haven, CT: Yale University Press.
- Moore, Jennifer. 1990. "What Is Really Unethical about Insider Trading?" *Journal of Business Ethics* 9: 171–82.
- Parkin, Michael. 2011. *Economics*. 10th edn. Boston, MA: Pearson Addison-Wesley.
- Pettifor, Ann. 2014. *Just Money: How Society Can Break the Despotism of Finance*. Margate: Commonwealth Publishing.
- Pilbeam, Keith. 2010. *Finance & Financial Markets*. 3rd edn. Basingstoke: Palgrave Macmillan.
- Plato. 380 BCE. *Republic*. Robin Waterfield (trans.). Oxford: Oxford University Press (2008 edition).
- Priyadarshree, Anurag, and Asad Ghalib. 2012. "Over-Indebtedness, Coercion, and Default: Causes of the Andhra Pradesh Microfinance Crisis and Regulatory Implications." *Enterprise Development and Microfinance* 23: 185–200.

- Ragatz, Julie, and Ronald F. Duska. 2010. "Financial Codes of Ethics." In John R. Boatright (ed.), *Finance Ethics: Critical Issues in Theory and Practice* (Hoboken, NJ: John Wiley & Sons), pp. 297–323.
- Reinhart, Carmen M, and Kenneth S. Rogoff. 2009. *This Time Is Different: Eight Centuries of Financial Folly*. Princeton, NJ: Princeton University Press.
- Richardson, Benjamin J., and Wesley Cragg. 2010. "Being Virtuous and Prosperous: SRI's Conflicting Goals." *Journal of Business Ethics* 92: 21–39.
- Roodman, David. 2012. *Due Diligence: An Impertinent Inquiry into Microfinance*. Washington DC: Center for Global Development.
- Rosenberg, Richard, Adrian Gonzalez, and Sushma Narain. 2009. "The New Moneylenders: Are the Poor Being Exploited by High Microcredit Interest Rates?" *CGAP Occasional Paper* 15. Washington DC: CGAP.
- Rothbard, Murray. 1983. *The Mystery of Banking*. New York: Richardson & Snyder.
- Ryan, John A. 1902. "The Ethics of Speculation." *International Journal of Ethics* 12: 335–47.
- Sandberg, Joakim. 2008. *The Ethics of Investing. Making Money or Making a Difference?* Gothenburg: Acta Universitatis Gothoburgensis.
- Sandberg, Joakim. 2011. "Changing the World through Shareholder Activism?" *Nordic Journal of Applied Ethics* 5: 51–78.
- Sandberg, Joakim. 2012. "Mega-Interest on Microcredit: Are Lenders Exploiting the Poor?" *Journal of Applied Philosophy* 29: 169–85.
- Sandberg, Joakim. 2013. "(Re-)Interpreting Fiduciary Duty to Justify Socially Responsible Investment for Pension Funds?" *Corporate Governance: An International Review* 21: 436–46.
- Sandberg, Joakim. 2016. "Pension Funds, Future Generations, and Fiduciary Duty." In Iñigo González-Ricoy and Axel Gosseries (eds.), *Institutions for Future Generations* (Oxford: Oxford University Press), pp. 385–99.
- Schlichter, Detlev. 2014. *Paper Money Collapse: The Folly of Elastic Money*. 2nd edn. Hoboken, NJ: John Wiley & Sons.
- Schumpeter, Joseph A. 1954. *History of Economic Analysis*. London: Allen & Unwin.
- Searle, John. 1995. *The Construction of Social Reality*. New York: Free Press.
- Shiller, Robert. 2012. *Finance and the Good Society*. Princeton, NJ: Princeton University Press.
- Sinn, Hans-Werner. 2010. *Casino Capitalism: How the Financial Crisis Came About and What Needs to Be Done Now*. Oxford: Oxford University Press.
- Smaga, Pawel. 2014. "The Concept of Systemic Risk." SRC Special Paper No 5, Systemic Risk Center.
- Smith, Adam, 1776. *An Inquiry into the Nature and Causes of the Wealth of Nations*. London: W. Strahan and T. Cadell.
- Sorell, Tom. 2015. "Is There a Human Right to Microfinance?" In Tom Sorell and Luis Cabrera (eds.), *Microfinance, Rights and Global Justice* (Cambridge: Cambridge University Press), pp. 27–46.
- Sorell, Tom, and John Hendry. 1994. *Business Ethics*. Oxford: Butterworth-Heinemann.
- Soros, George. 1987. *The Alchemy of Finance*. Hoboken, NJ: John Wiley & Sons.
- Soros, George. 2008. *The Crash of 2008 and What It Means*. New York: Public Affairs.
- Strange, Susan. 1986. *Casino Capitalism*. Oxford: Basil Blackwell.
- Strudler, Alan. 2009. "Insider Trading: A Moral Problem." *Philosophy & Public Policy Quarterly* 29 (3/4): 12–16.
- Thamotheram, Raj, and Aidan Ward. 2014. "Whose Risk Counts?" In James P. Hawley, Andreas G.F. Hoepner, Keith L. Johnson, Joakim Sandberg, and Edward J. Waitzer (eds.),

- Cambridge Handbook of Institutional Investment and Fiduciary Duty* (Cambridge: Cambridge University Press), pp. 207–21.
- Thomas, Abulkader (ed.). 2006. *Interest in Islamic Economics: Understanding Riba*. New York: Routledge.
- Visser, Wayne A.M., and Alastair MacIntosh. 1998. “A Short Review of the Historical Critique of Usury.” *Accounting, Business and Financial History* 8: 175–89.
- Walsh, Adrian, and Tony Lynch. 2008. *The Morality of Money*. Basingstoke: Palgrave Macmillan.
- Warde, Ibrahim. 2010. *Islamic Finance in the Global Economy*. Edinburgh: Edinburgh University Press.
- Weatherford, Jack. 1998. *The History of Money*. New York: Three Rivers Press.
- Weber, Heloise. 2004. “The ‘New Economy’ and Social Risk: Banking on the Poor?” *Review of International Political Economy* 11: 356–86.
- Weber, Max. 1905. “Die Protestantische Ethik und der Geist des Kapitalismus.” *Archiv für Sozialwissenschaften und Sozialpolitik* 21: 1–110.
- Werhane, Patricia H. 1989. “The Ethics of Insider Trading.” *Journal of Business Ethics* 8: 841–45.
- Werhane, Patricia H. 1991. “The Indefensibility of Insider Trading.” *Journal of Business Ethics* 10: 729–31.
- Wesley, John, 1960. “The Use of Money.” In *Sermon on Several Occasions*, Sermon 50. London: Epworth Press.
- Yunus, Muhammad. 1998. *Banker to the Poor*. Dhaka: University Press Limited.
- Yunus, Muhammad. 2007. *Creating a World without Poverty: Social Business and the Future of Capitalism*. New York: PublicAffairs.

CHAPTER 18

ETHICS AND, IN, AND FOR LABOR MARKETS

MICHAEL S. McPHERSON AND DEBRA SATZ

18.1 INTRODUCTION

THE classical political economists noted three important features of the markets in which human labor power is bought and sold. First, such markets can have endogenous effects on the very workers whose labor power is being sold. Consider Adam Smith's remark on the worker in the pin factory:

the understandings of men are necessarily formed by their ordinary employments. The man whose whole life is spent performing a few simple operations of which the effects too are perhaps always the same... has no occasion to exert his understanding or to exercise his invention in finding out expedients for removing difficulties which never occur. He naturally loses, therefore, the habit of such exertion and generally becomes as stupid and ignorant as it is possible for a human creature to become... of the great and extensive interests of his country, he is altogether incapable of judging.

(1776: 781–2)

From a different perspective, Marx (1867) worried that in a capitalist society workers would be reduced to mere appendages to machines. Because labor power is embodied in human beings, who may be deeply affected by the conditions under which their labor is sold, the democratic state has important regulatory interests with respect to these markets. Democratic states depend on having informed citizens whose members relate to one another as equals (Marshall 1950). If the labor market tends to produce people who are unable to function as democratic citizens, then the labor market needs to be changed or institutions designed outside the labor market that can counteract its effects.

Second, labor markets involve exchanges whose terms cannot be completely specified and written down. It is impossible to determine precisely when and whether a worker is

working as hard as she can. Both the quality and the intensity of work cannot be fully specified in a contract. This means that the contracting parties have to rely on trust, as well as other norms in instantiating the terms of their agreement. Moreover, although some of their interests may overlap, owners and workers also have conflicting interests. Owners are worried that workers will shirk, whereas workers are worried that owners will use their power to exploit them.

And third, the parties to the exchange—workers and owners—have an asymmetric ability to sanction one another. Because work effort cannot be fully specified, employers have an interest in finding ways to induce workers to work hard. A worker has no incentive to work hard when the results of her effort simply go to her boss (and not to her own pay). Similarly, a worker may have no incentive to work hard when, if fired, she can easily find another job at almost no cost. These two last features of labor markets show that it can be rational for employers to pay their workers more than the equilibrium wage in order to be able (because such markets will now not clear in equilibrium) to use the threat of unemployment to discipline workers (Bowles and Gintis 1990).

The fact that labor markets have endogenous effects, that labor contracts are incomplete, and that workers and owners have some conflicting interests, distinguishes them from other types of markets such as markets where wheat, oil, apples, or oranges are being exchanged. Not only do markets usually not have important feedback effects on these latter products, but also a buyer can specify more precisely the amount and grade of wheat or oil he wants. And in the case of asymmetric power, we need not worry about the mistreatment of wheat and oil; what happens to the human beings who produce the wheat and oil, however, is always a cause of concern.

In this chapter, we explore the consequences of the special nature of labor markets for three specific issues: unemployment, the rise of the gig economy, and the nature and organization of work in a just society. Each of these issues involves the complex interplay of economic considerations with ethics.

18.2 LABOR MARKETS, LEGITIMATE WAGES, AND INVOLUNTARY UNEMPLOYMENT

If labor is a good like any other, then we should expect that the law of supply and demand would determine its price. One way to understand this is that the price that any laborer can command on the market is equal to the value of her marginal product, the price she can demand for forgoing a marginal hour of leisure. If workers do not find work on this assumption, it must be because they prefer leisure to the wage that the value of their marginal product would set for them in the market. We might think of it this way: unless one's labor is of absolutely no productive value whatsoever, there is some wage—perhaps very low—that a worker can command on the market. If she declines to work for that wage, it must be because she prefers not working. Hence, all unemployment must be voluntary.

Making an argument along these lines, Nobel Laureate Robert Lucas writes:

The unemployed worker at any time can find some job at once and a firm can always fill a vacancy instantaneously. That neither typically does so by choice is not difficult to understand given the quality of the jobs and the employees which are easiest to find. Thus there is an involuntary element in all unemployment, in the sense that no one ever chooses bad luck over good; there is also a voluntary element in all unemployment, in the sense that however miserable one's current work options, one can always choose to accept them. (1978: 354)

Relying on some standard economic assumptions, it is indeed hard to see how completely involuntary unemployment is possible. Markets will clear unemployment; if there is an excess supply of labor then the price of labor should fall. In that case, unemployment in the market will exist only insofar as workers without jobs are making a rational choice not to work on the basis of all the information and alternatives available to them. For this reason, many economists have also argued against regulations—such as minimum wage laws and rules governing unionization and work—all of which they believe discourage employers from hiring more workers and interfere with the ability of the labor market to clear. On this view, if we remove the restrictions to markets, then involuntary unemployment will be impossible.

By contrast, some who argue that involuntary unemployment exists claim that some workers must be off the labor supply curve. For example, as we noted above, labor markets will not clear if employers pay workers more than the market wage to enhance their productivity. If labor markets cleared and exit was costless to the worker, the worker might be tempted to shirk. So firms have a reason to give employees an incentive to remain at work. If all firms behave in this way, then there will be some workers who cannot find work, but who would be willing to work at prevailing wages.

Keynes (1937), by contrast, explained involuntary unemployment, in part, on the basis of inadequate aggregate demand for goods and services. It is worth rehearsing his argument here. Keynes claimed that lowering wages to allow the labor market to clear would, contrary to the arguments of many economists, not actually reduce unemployment. Rather, the cut in wages, by reducing employee income, would decrease consumer spending. This in turn would reduce the demand for products, leading to a further reduction in production and investment, forcing companies to not only cut wages, but lay off employees. Now the employees (and former employees) are even more unable to spend, decreasing both consumer spending and demand. According to Keynes then, the real wage is not co-determined with the level of unemployment; rather unemployment is a result of a fall in the level of aggregate demand. As he put it,

the volume of employment does not depend on real wages, except in so far as the supply of labour available at a given real wage sets a maximum level of employment. The propensity to consume and the rate of new investment determine the volume of employment, and the volume of employment determines the real wage. (1973: 377)

Keynes believed that the government needed to intervene to raise the demand for workers by stimulating demand through monetary or fiscal policy.

A fourth explanation for why employers do not lower their wages in the face of unemployment is that the unemployed workers are not perfect substitutes for the currently employed workers. The unemployed are less prepared to work efficiently. They might be willing to work for a lower wage, but not for a wage that is low enough, once we take training and transaction costs into account (Varian 1976).

It is tempting to see the debate over involuntary unemployment as an entirely empirical matter, focusing on, for example, the question of whether labor markets will clear in equilibrium. But that is not the case: regardless of whether labor markets will clear, we can still ask whether the choices that are available in such markets should be considered voluntary. Even if workers are rationally making choices to accept or reject employment at drastically lowered wages, that does not mean that such choices are voluntary choices.

Many philosophers would argue that the concept of “voluntary” has a separate normative dimension. Consider a person’s response to the robber who offers “your money or your life.” It is certainly rational for the person in that case to part with her money. But is the choice truly voluntary? In one sense, it may be: the person on her own makes the appropriate movements to hand over her money to the robber. But in another, it is not: the robber has no entitlement to her money and the two choices on offer are therefore illegitimate.

On this understanding, choices count as “voluntary” only with respect to a moralized baseline. Alan Wertheimer (1990) argues that choices are voluntary only when (1) there are acceptable alternatives and (2) there is no violation of a right. According to Wertheimer’s moralized account of voluntariness, “your money or your life” is a canonical example of coercion, not an offer of a free, voluntary choice. With respect to unemployment, on this view, the choice to be unemployed is only voluntary if the work on offer is “acceptable,” and if the terms of work do not constitute a violation of the worker’s rights.

There are other options for understanding the actions involved in the robber example. Some philosophers grant that the victim’s actions are voluntary, but they also deny that the voluntary choice in this question (“hand over the money”) makes the outcome legitimate. Similarly, someone might deny that his voluntary choice involved in rejecting a wage offer is sufficient, on its own, to justify his being unemployed. The wage offer might be indecent.

Of course, many people *assume* that showing that a choice was voluntary is sufficient to justify it. And that is why how we frame the question of choice can seem so important: conservative economists argue that the unemployed simply lack the willingness to make an effort, and liberal economists claim that people are largely unemployed through no fault of their own. Consider, as a representative example of this association of ideas, this passage from Herbert Spencer:

On hailing a cab in a London street, it is surprising how frequently the door is officiously opened by one who expects to get something for his trouble. The surprise lessens after counting the many loungers about tavern-doors, or after observing the

quickness with which a street-performance, or procession, draws from neighbouring slums and stable-yards a group of idlers. Seeing how numerous they are in every small area, it becomes manifest that tens of thousands of such swarm through London. “They have no work,” you say. Say rather that they either refuse work or quickly turn themselves out of it. They are simply good-for-nothings, who in one way or other live on the good-for-somethings. (1884: 32)

Spencer argues that the poor are idle because that is what they have chosen to do; since they are responsible for having chosen their plight, they are not entitled to our aid. But looking at the issue through the window of ethics shows us that what is really at stake in the debate over unemployment is, in the main, our view of the prior entitlements we think poor people have. In a sense, then, thinking of coercion and voluntary choice as a freestanding moral problem just distracts us from the central issue at stake in debates about unemployment: do workers have an entitlement to jobs at an “adequate” standard of living? Notice that if workers were apples or cars, the issue of entitlements would not arise. Apples are not entitled to anything.

Other ethical issues surround the question of unemployment beyond the question of whether or not it is voluntary. For example, *who* exactly should be included in the calculation of the unemployment rate? Many critics believe that current methods of measuring unemployment are inaccurate. They do not count the 1.5 percent of the available working population incarcerated in US prisons, those who are no longer actively looking for work, those who are self-employed, and those who increasingly work part-time but would like to work full-time. And of course, there is the question of how to deal with those who come in and out of the labor market over their working lives: mothers of young children, recent grandparents, and so on. Surrounding these issues is the question of whether our standard model of unemployment is applicable to current conditions where “non-standard” work seems on the rise. This brings us to our second topic.

18.3 LABOR MARKETS AND THE RISE OF THE GIG ECONOMY

Over the decade from 2005 to 2015, the share of US workers in “nonstandard” employment relationships expanded from 10 to 16 percent of the workforce. (Katz and Krueger 2016: 1). The “nonstandard” category includes persons who find work through temp agencies and persons who supply “contracted labor” for companies that have outsourced functions like janitorial services, as well as people who work “on call” (holding themselves available for work) or who get temporary jobs through union hiring halls. The category also includes self-employed people who work as consultants, rather than through traditional employment.

Among these “contractors” are participants in the “gig economy.” Uber drivers, for example, are in most cases not employed by Uber, but instead are viewed as working on

contracts with their passengers, with Uber serving as a “match maker” or intermediary. Such internet-mediated services include not only ride sharing but also general task performance (such as Mechanical Turk and TaskRabbit), restaurant food delivery (Grubhub), and grocery shopping (InstaKart), among others.¹ The gig economy did not even exist in 2005 and while it remains a small fraction of nonstandard employment (probably less than 10 percent), it is growing quite rapidly and in the view of some observers is destined to become a major factor in the US economy.

Remarkably enough, *all* of the net growth in employment in the United States over the 2005–15 decade was found among these various categories of nonstandard or contingent work; traditional employment did not grow in the aggregate at all. While a variety of factors, including the Great Recession, have probably contributed to the rapid growth in contingent and contracted employment, a common characteristic of all these forms of work relations is that they attenuate the connection between workers and employers or firms and in many cases leave workers with fewer rights than traditional employees enjoy. If this is indeed a lasting trend, it may well have worrisome implications for the status of the American work force.

Existing labor law assures both wage and salary workers of a set of rights and regulatory protections that are not available to contractors. These include minimum wage regulations, unemployment insurance, the right to bargain collectively, extra pay for overtime, and regulations that protect workers from arbitrary dismissal. These protections can be expensive for firms to provide (although there may be some compensation in terms of a better motivated and more cooperative workforce).

The problem in labor law of defining the boundary between what counts as an employment relationship versus a purchase of services agreement is not new, and in many cases the interests of workers and employers conflict as to the precise location of it. In general, the defining characteristics of contractors have been taken to include providing their own tools, having freedom to set their own hours of work, and having some discretion over how the work is performed. Businesses are sometimes counseled not to provide contractors with facilities like an office or an email address that may be seen as implying a traditional employment relationship, or providing a company-supplied computer that may compromise their independence. In many cases, taxi drivers do not own the vehicles they drive, which helps explain why taxi drivers are often considered employees while Uber drivers, who use their own cars, are not.

These issues are arising in a particularly “modern” form with the swift emergence of the gig economy. Gig economy businesses grew up in the same Silicon Valley soil as personal computing devices, social media, and search engines. It is no surprise that the people who fashioned devices and software that could transfer enormous amounts of information among people with great speed and little fuss about ownership should embrace slogans like “information wants to be free,” should see governmental restrictions on these flows with suspicion, and should be congenial to a libertarian outlook

¹ The gig economy also includes enterprises that coordinate room or house rentals (like AirBnB) or other physical property. We are keeping the focus here on the particular matter of selling one’s labor.

more generally. At least on the surface, the transactions that pervade the internet are quintessentially Robert Nozick's "capitalist acts between consenting adults" (1974: 163).²

Uber, for example, has tried to market itself as merely a platform that facilitates a private transaction between a driver and passenger. Uber's role, so they say, is simply a modern version of two people on a telephone arranging a ride to the airport. Task Rabbit, InstaKart, and others might similarly claim that they are simply creating a market in which people can work out their own agreements.

Uber and its business analogues, however, play a vital role in standardizing the terms of the arrangements clients and suppliers—riders and drivers—make. If Uber were purely a silent partner, simply introducing drivers and passengers to one another, the ensuing negotiation over price and other terms would make costs prohibitive. Instead, Uber imposes a quite substantial regulatory structure governing deals between riders and drivers. (Although details differ, similar structures exist for other gig economy enterprises.) Much as cab companies, often guided by municipal regulators, set standardized fares, Uber provides the service of lowering transactions costs by establishing a fare structure and imposing rules to block drivers and riders from negotiating around it. Uber standardizes these transactions in ways that go well beyond price. Without regulation by Uber, it would be costly for riders to assure themselves that the unknown driver had a safe, reasonably clean, and up-to-date car, while at the same time releasing the driver from the obligation and risk involved in collecting the fare herself. And both rider and driver are protected by knowing that Uber has lots of information about both parties should anyone misbehave.

Moreover, and in contrast to Nozick's formulation quoted above, other people besides the passenger and driver may have a stake in this transaction. The claim that transactions between Uber drivers and passengers are purely private ignores possible externalities that affect other people. For example, the risk of unsafe Uber cars and drivers threatens the safety and well-being of other drivers and passengers in other cars. An additional risk is that unknown drivers may not have insurance that applies when their vehicle is used for commercial purposes. Because Uber can monitor the driving records of drivers, the safety of the cars they drive, and the insurance they carry (or which Uber carries on their behalf) far more effectively than riders can, the company has a moral responsibility, and arguably should have a legal responsibility, to attend to these concerns.

Some observers have further asserted that Uber and similar ride services (like Lyft) add to pollution levels and congestion in busy areas, both classic examples of externalities. Whether these external effects are significant is a difficult empirical question.³ If Uber and its kin substitute for taxis or discourage people from owning or driving their own cars, their presence may have no net impact or could even reduce pollution and crowding. But if Uber instead substitutes for mass transit, its negative effects could be

² In fact, no person or business wants all information to be free. Persons generally value their privacy and businesses often depend for their profits on patents and trade secrets. Thus the owners of Google, while generally passionate about making information flows free, have never been inclined to publish the details of their search algorithms.

³ See O'Donovan and Singer-Vine (2016).

non-trivial. These effects could justify enforceable regulations to limit the number of taxis and ride-sharing vehicles on the road at any one time (as New York City's medalion system has done in the taxi industry).

A distinctive feature of gig economy businesses is their potentially disruptive effects on the application of labor law. As we noted above, if this type of work arrangement continues to expand, the treatment of these work arrangements may have real consequences for the labor market as a whole. As we noted earlier, the enterprises in the gig economy that connect clients with providers of services have claimed that the service providers are contractors, not employees, and this interpretation has generally prevailed so far. This is a consequential matter for companies like Uber, since their profits would be considerably lowered if they had to treat their drivers as benefited employees.⁴ If it does turn out that gig economy jobs continue to expand rapidly as a fraction of total employment, then this matter also becomes consequential for the economy and the workforce as a whole.

A particularly revealing question to consider is the applicability of minimum wage laws to gig economy workers. Consider this hypothetical: Imagine someone, call her Ms. A, who has a full-time day job that pays above the minimum wage but wants to earn some extra money saving up for a Caribbean vacation. She has a car that meets Uber's standards and she decides to drive on evenings and weekends to get some extra cash. She knows she might turn out to earn less than the federal minimum hourly pay (after allowing for the fuel and maintenance costs on her car), but it is an easy option for her to try and she can always quit if it does not work out.

Putting aside the broader matter of externalities noted above, why is her decision to work at a below-minimum wage anybody's business but her own, if this is how she wants to spend her time? Viewed as an isolated transaction, it is indeed hard to see the harm. But there is an entire class of transactions into which this example fits: people with full-time jobs who drive part-time to supplement their incomes. This class fits into a larger class of people who work part time to supplement a full-time job, perhaps by packing groceries, and the still larger class of people who work part-time at a job, regardless of their other employment.

If Ms. A should be allowed to waive the minimum wage to work part-time, we then need to ask how her situation differs from that of other part-time workers, suggesting that perhaps the minimum wage should only apply to full-time workers. The problem is such a rule would tend to be destabilizing. Employers would have an incentive to organize production so as to make as much work as possible part-time, thus creating opportunities to pay less than the minimum wage.

Other exceptions to the rule that employers should not pay workers less than the minimum wage (and that workers should not be allowed to accept such work) tend to be similarly destabilizing. Employers will have an incentive to provide employment in a form that fits the exemption, regardless of whether that form of organizing employment is actually more productive than others. The implication is that if we want to have a

⁴ A recent court decision in the influential state of California may make it more difficult for companies to impose contractor status on gig economy workers; see Khouri and Lien (2018).

minimum wage at all, it is valuable to make it as nearly universal as possible. If gig economy workers are exempt from the minimum wage, as they are when they are classified as contractors, then employers will prefer to deliver services in gig economy form, and the economy will be distorted thereby. Indeed, firms are increasingly choosing to “buy” from outside contractors much of what they used to “make” internally using their own employees. This is what David Weil (2014) calls the “fissuring” of work, and it includes everything from the McDonald’s franchise model, to Uber’s “we’re just the platform” model, to Apple’s outsourcing of iPhone and iPad production to Foxconn. As long as the gig economy is small relative to the labor market, the collective effect may be trivial in practice, but the future shape of the economy, and the future status of workers, may be influenced by regulatory decisions made early on.

If we want a minimum wage to prevail, there are then good reasons to make it as universal as possible. And this invites the question: why have a minimum wage at all? There are at least two answers, one largely pragmatic and the second more principled. The first answer is grounded in a judgment that the distribution of income in our society unfairly advantages owners relative to workers and the minimum wage tends to raise workers’ wages. It is one tool that governments can use to attempt to redistribute income from employers toward workers. This mechanism applies straightforwardly to those who would be paid less than the minimum if that were legal, but a minimum wage may also raise the wages of those who make more than the minimum by strengthening their bargaining power.⁵

The second argument is grounded in a conception of the dignity of labor, and we alluded to something like it in our earlier discussion of involuntary unemployment. Demanding of some people that they accept work at pay below a reasonable minimum undermines their status as equal citizens. Jobs that are economically viable only at wages below a minimum set by society should not be done at all. Thus there are both practical and principled reasons why minimum wage laws should apply to workers in the gig economy, if doing so is feasible.

In their very instructive recent paper on the gig economy, Harris and Krueger (2015) find the question of minimum wage feasibility for gig workers a challenging one. A driver’s time waiting in her car for a call is presumably work time, but she may be waiting for a call from Lyft and from Uber at the same time, in which case it is hard to say for whom she is working, except possibly herself. Moreover, drivers are not required to accept every ride they are offered, and being “fussy” about selecting rides can be seen as a voluntary choice to lower your wage. Even worse measurement problems may attend other types of gig economy jobs in which workers bid to perform particular tasks (such as TaskRabbit and Mechanical Turk). Harris and Krueger wind up concluding that while many protections should apply to gig workers, the minimum wage is not among them.

⁵ How powerful the minimum wage as a redistributive strategy is controversial. Economists have long recognized that a higher minimum wage may reduce employment, perhaps by enough to offset the benefit of a higher wage for those who keep their jobs.

Although Harris and Krueger point to real practical difficulties, they are generally alive to the concern that expanding the scope of contract workers and reducing traditional employment may result in economic distortion and worsen conditions for labor. They recognize that the development of the gig economy—and more generally the move from traditional employment to contracted labor—threatens the social compact reflected in institutions like unions, collective bargaining, and wage and hours regulation. They believe that decisions about whether a company should organize around gig economy relationships or more traditional employment should be based on their relative economic efficiency and not on the ability of the former arrangement to keep worker pay and benefits relatively low. They call this the principle of “neutrality.”

To help maintain neutrality, Harris and Krueger propose defining in the law a third class of workers between standard employees and independent contractors. They call this new class “independent workers,” and suggest that in some important respects these workers should receive the same protections granted to standard employees, including the right to unionize and to bargain collectively. But they claim that the difficulty in measuring hours of work in gig economy jobs blocks them from being protected by the minimum wage and overtime rules.

The measurement difficulties are real, but the costs of denying minimum wage protection to gig employees may also be substantial. Harris and Krueger argue that the absence of minimum wage regulation may not be so serious, because collective bargaining rights will help protect workers from low wages, and also because the low cost of entering and exiting the market for gig economy work helps justify the claim that lower gig economy pay is a fair trade off for greater flexibility. They view the decision to work for less as essentially voluntary.

This argument seems worryingly context dependent. Right now gig economy workers are a tiny fraction of the labor force and traditional employment—especially in an economy as strong as the US is experiencing as of this writing—is relatively easy to obtain. However, many observers expect the share of jobs that are mediated through internet-based connections between customers and suppliers to grow rapidly, and the demand for such services will only accelerate if low wages are allowed to prevail.

This point seems to argue in favor of providing for an enforceable minimum wage. Just how impossible is it to measure the working hours of such workers? At least in the case of Uber drivers, the fact that several recent papers have reported estimates of the pay of Uber drivers suggests there are encouraging prospects for effective measurement here.⁶ More fundamentally, the concern here is to have regulations that ensure that a big loophole is not being introduced into the minimum wage law. Our major concern is to guard against a substantial fraction of workers operating in circumstances where the minimum wage is clearly out of reach. Thus one might want to address a question

⁶ These of course are measures of average wages in some samples. In those samples reported on, average drivers earned well above the current federal minimum of \$7.25 per hour. Of course, averages don't speak directly to the question of whether some workers receive less than the minimum. See, for example, O'Donovan and Singer-Vine (2016).

like the following: consider a reasonably conscientious driver who worked for several hours in reasonably favorable conditions (that is, not too much down time between fares). What is the probability that she would earn less than the minimum wage? A threshold probability might be set for purposes of enforcing minimum wage regulations. Admittedly, this is a lot harder and less precise than measuring the minimum wage in traditional employment, but it may be worthwhile if it guards against widespread violation of Harris and Krueger's principle of neutrality—a principle whose violation itself is difficult to discern. There are other types of work arrangements where the Department of Labor has had to develop special rules to safeguard against loopholes, including minimum wage for tipped workers and for those who perform piecework in their homes.

A key point here is that the logic of the minimum wage (and other labor regulations) is one of collective action: the larger the fraction of workers operating under the minimum wage, the more effective and stable the level of protection is likely to be. As we noted earlier, the gig economy is only one, so far minor, example of the growth of contracted employment in place of traditional employment through growth of temporary workers, outsourcing to firms with contracted workers, and greater use of contractors. A number of factors, including the recent deep recession, may have contributed to this trend, but one result of the trend (and very likely a cause) is an attenuation in the relationship between companies and those who do the company's work. In the absence of effective regulation, it is likely that this trend will leave workers increasingly vulnerable to the vagaries of the market for their services.

Firms operating in the gig economy, like other businesses driven by internet technology, including Facebook, Amazon, and Google, generally benefit from network economies that reward scale. The result is a growing trend toward the forging of strong monopolies or oligopolies for these firms on both sides of the market, as both employers and service providers. This trend not only produces unfortunate economic inefficiencies but it can be highly disempowering both for consumers and workers. The gig economy exacerbates workers' sense of powerlessness, dooming them to a reverse auction in which they are constantly pressured to reduce their wage demands in order to out-compete rivals. It is hard to imagine successful collective action among independent contractors to demand better terms—especially when federal agencies are likely to view such actions as violations of antitrust laws. An important role of government throughout the twentieth century has been to countermand trends in business and technology that threaten the pay and status of workers, and that role is clearly important in the twenty-first century as well.

18.4 LABOR MARKETS IN A JUST SOCIETY

Our final topic involves the organization of work. Although most people in the developed world spend a majority of their days at work, there has been little attention paid by recent political philosophy to the organization of work, or the kinds of relationships that

workers and owners have with one another. Indeed, a default assumption running through much of recent liberal philosophy is that there is a large gap between the economic and political structures of society: while the political structure must be democratic, there is no parallel case for democratizing the workplace. Where theorists of right and left divide is over the extent to which the government is entitled to regulate the economy as a whole. For example, in *A Theory of Justice* (1971), Rawls develops his two principles which are to regulate the basic structure of society. Rejecting both libertarianism and state command economy, Rawls left open the nature of economic organization, an omission he later regretted. In *Justice as Fairness* (2001: 178), he writes with concern that the account of justice as fairness “has not considered the importance of democracy in the workplace and in shaping the general course of the economy.”

Those who defend hierarchical, non-democratic authority in the workplace generally do so on two grounds. The first is that if workers have the right to exit employment, that fact should be sufficient to curb the worst abuses. “Exit,” on this view, ensures the protection of workers’ interests: by threatening to leave, workers can rein in managerial and employer power (Taylor 2017). In response, we have already noted that where labor markets do not clear, there will be unemployment. Where there is unemployment, the right of exit may turn out to be more formal than substantive. Additionally, some workers may lack the skills and training, or the relevant information, to pursue other opportunities; employers’ use of non-compete agreements limits workers’ ability to exit from their jobs; and many workers cannot relocate for work but find themselves in environments with few employers. Furthermore, it might be argued that exit by itself does not entail that there are any democratic workplaces to join: perhaps all the workplaces on offer are hierarchical. This is likely to be the case because workers rarely have access to the capital needed to start a company of their own.

The second ground for defending hierarchy and managerial power is that these are said to be good for the efficiency of firms. Where contracts are incomplete, it is costly to determine and specify what is needed from workers in advance. Managers need discretion to coordinate worker’s activities, assign tasks or projects as they arise, and to oversee the work or alter its direction as needed. This “Coasean” theory of the firm (Coase 1960) helps explain why substantial managerial control makes sense.

It is important, however, not to caricature a democratic alternative to the hierarchical firm. A cooperative or employee-owned firm need not take the form of a direct democracy in which every significant decision is made by a vote of the employees as a whole. An alternative would be for the collective owners to designate one or more employees as managers, with authority to make decisions subject to rules established collectively and overseen by an elected subset of employees. This is, of course, roughly the way things work in a representative political democracy (Ellerman 1990).

Moreover, the assumption that privately owned firms will be efficient is vulnerable. The immediate incentive for the owners is to maximize their own income, which is not the same as maximizing the business’s overall contribution to the economy. For example, if the firm has monopsony power, it will reduce the size of the workforce and therefore the firm’s output below the economically optimal level, in order to reduce wages. Another

possibility is that workers will withhold information, for example, about shirking by other workers that they would report if they saw themselves as having a stake in the firm's success (McPherson 1983).

If workplace rules were only dictated by efficiency, there would still be problems with ignoring the structure of the workplace. There are serious issues about when to allow efficiency to outweigh other goals we might have for a productive enterprise; workers are people rather than things and we have reason to care about their well-being and safety, even if it were efficient to dispense with them. But workplace rules are not merely designed to enhance efficiency. In many workplaces, employers regulate and control how their employees dress and their modes of speech, and they can sanction workers for whom they marry, what causes and political parties they support, and even how they wear their hair. Although there may be idiosyncratic cases where one or another of these restrictions has an efficiency rationale, the more plausible explanation for most such rules is simply owners expressing their own preferences and then imposing these preferences on their workers. Additionally, US law allows employers to terminate workers at will, although within some important limits: for example, employers cannot discriminate against protected categories and they cannot engage in retaliation with respect to protected activities. Thus, employers can fire workers whose dress style or cultural affiliation they do not like. They can alter the terms of their work, such as their salaries. And even in the absence of explicit employer actions, the threat of termination may be enough to lead many workers to toe the company line. Unions, once a countervailing force to such asymmetric power, now constitute only about 7 percent of the private employer workforce.

In her recent Tanner lectures, Elizabeth Anderson (2017) argues that Americans tolerate a degree of dictatorship in the workplace that they would never accept in their government. She notes that some large employers search through the personal belongings of their employees, prevent their workers from using the bathroom during work hours, and that about half of all US employees are subject to unmotivated drug testing by their employers. Many such displays of managerial power are episodic and arbitrary and extend to workers' behavior even when not at work.

There are many arguments given in the history of political thought for pressing the case for workplace democracy. We focus on two. The first argument, associated with the writings of John Stuart Mill (1848), is that hierarchical workplaces are incompatible with the autonomy of free citizens. Hierarchical workplaces dull the habits of free inquiry, critical reflection, and treat citizens as mere dependents who cannot exercise choice. Workers schooled in such workplaces, Mill believed, would fail to develop the capacities that they needed to deliberate as democratic citizens. His argument looks to workplace democracy as a crucial arena to provide "psychological support" for political democracy. Also, as we saw earlier, in *The Wealth of Nations* (1776) Adam Smith had similarly worried about the political effects of a division of labor which left little room for human reflection on tasks and even less room for critical thought.

A second argument contends that the same reasons that justify democracy in the political realm also justify democracy in the economic realm. Michael Walzer (1978)

recounts the story of George Pullman, who founded and autocratically ruled a town where his employees lived, a “company town.” Walzer claims that if we accept the premise that just because Pullman founds a town it does not entail that he owns that town and can make any laws he wishes, then we should also accept that his founding the company does not justify his having full ownership of that company for which he can make any rules that he wants. (The Illinois Supreme Court denied owners full control rights over the towns they built to house their workers.) The fact that Pullman took some risks in establishing the town does not grant him the rights of a feudal overlord. Instead, the people who are most affected by the town’s policies—the residents who live there—ought to have the right to govern it. Walzer’s argument is that there is a “parallel case” between the town and the company. The risks owners take in founding a company do not give them the right to determine everything about the company. In both cases, what affects all, Walzer claims, should be decided by all, independent of how these structures came about. In both cases democratic decision-making should prevail.

Anderson’s own analysis highlights the abuses that occur when power in any realm is unchecked. To be sure, unlike Walzer and Mill, she does not embrace robustly democratic workplaces. She accepts (perhaps too quickly) that such workplaces would be inefficient; it is certainly true that they would be difficult to achieve. Instead she argues that we urgently need a reform agenda that includes new worker rights, including more extensive protections of workers’ freedoms off the job, such as their rights of free speech, their rights to organize, and their freedom to exercise sexual choice—and increased worker participation in firm governance. While she does not specify the form that the latter representation of worker’s interests should take, she clearly has in mind the need for some mechanisms to ensure that workers’ views are heard, fairly considered, and have influence. As an example, firms in some European countries have work councils, consultative bodies where all the workers in the firm are represented in decision-making. Worker representatives also can sit on company boards; unions can demand more of a say in company decisions such as whether or not to outsource or relocate. Going beyond such measures, labor organizations could lobby for credit subsidies to make worker owned enterprises more of a real possibility.

Examples of the encroachments of employers on the freedoms of their employees, such as those listed above, provide compelling evidence that there is a need for greater worker protections. In some cases, the legal basis for these protections exists but the protections are simply unenforced; in other cases, it is a question of creating new rights. Nonetheless, an account that focuses on the authoritarian power that firm managers exercise over their employees runs the risk of overlooking something important about workers’ vulnerability in the contemporary economy. As our previous section makes clear, a great deal of employer power is now exercised in a context where many employers have forgone direct managerial control. Uber simply does not care about the dress, hairstyles, or marriage partners of those who drive for them. In the context of the gig economy, and the growing outsourcing of work beyond the firm, rights *within* the workplace may turn out to depend on rights *outside* the workplace. Among such broader rights might be included social protections such as a decent minimum wage, job training programs, universal health care, and perhaps a universal basic income. Ironically, it might be argued

that workers' vulnerability comes less these days from what Anderson calls "private government" operating in the workplace and more from employers' withdrawal from the lives of their workers, and moves to turn more of their employees into the role of occasional "contractors." If so, a better solution to worker's disempowerment and abuse might be to enhance democratic control of the economy as a whole as well as to strengthen the bargaining position of all workers.

Anderson's call to expand the rights of workers more generally is apropos here. US labor law currently prohibits certain strikes and boycotts (such as secondary boycotts) and it restricts the basis for other strikes. Current antitrust law also prohibits organizing by independent contractors, a category that as we have seen, encompasses growing numbers of workers. Additionally, unions have no presumptive right to strike over decisions about the basic direction being taken by the firm or enterprise that employs them, including how and where the firm invests its money. Widespread and coordinated strike efforts by workers are also potentially vulnerable to antitrust legislation. Changing these laws would help redress the highly unequal power of workers and employers. Other measures that can strengthen workers' bargaining position include a national full employment commitment, a national minimum wage set at a living wage, and measures that aim to help displaced workers with job training and skill expansion. Also, countries like the United States and India might move away from decentralized wage bargaining toward a more coordinated model as is found in the Nordic countries. While these measures do not aim at workplace democracy per se, they do increase labor's bargaining position.

18.5 CONCLUSION

Whatever measures are ultimately endorsed in our changing economy, we hope to have shown that work and its organization involve questions not only of efficiency and power, but also fairness, equity, and autonomy. A recent Gallup poll reported that about three quarters of workers in the United States are "disengaged" at work, which not only costs employers in lost productivity, but also and most importantly takes a toll on workers' well-being (Harter 2017). Political philosophers need to address the ethical issues involved with the organization of workplaces, the nature of work, and the direction and control of the economy. These are matters that should not be addressed through economic considerations alone.

REFERENCES

- Anderson, Elizabeth. 2017. *Private Government: How Employers Rule Our Lives (And Why We Don't Talk About It)*. Princeton, NJ: Princeton University Press.
- Bowles, Samuel, and Herbert Gintis. 1990. "Contested Exchange: New Microfoundations for the Political Economy of Capitalism." *Politics and Society* 18: 165–222.
- Coase, Ronald. 1960. "The Problem of Social Cost." *Journal of Law and Economics* 3: 1–30.

- Ellerman, David. 1990. *The Democratic Worker Owned Firm*. London: Unwin Hyman Limited.
- Harris, Seth D., and Alan B. Krueger. 2015. "A Proposal for Modernizing Labor Laws for Twenty-First Century Work: The 'Independent Worker.'" *The Hamilton Project*, Discussion Paper 2015-10. Available at http://www.hamiltonproject.org/assets/files/modernizing_labor_laws_for_twenty_first_century_work_krueger_harris.pdf.
- Harter, Jim. 2017. "Dismal Employee Engagement Is a Sign of Global Mismanagement." *Gallup Blog*. December 20. Available at <https://news.gallup.com/opinion/gallup/224012/dismal-employee-engagement-sign-global-mismanagement.aspx>.
- Katz, Lawrence, and Alan Krueger. 2016. "The Rise and Nature of Alternative Work Arrangements in the United States, 1995-2015." NBER Working Paper No. 22667. Cambridge, MA: National Bureau of Economic Research.
- Keynes, John Maynard. 1937. "The General Theory of Employment." *Quarterly Journal of Economics* 51: 209-23.
- Keynes, John Maynard. 1973. *The Collected Writings of John Maynard Keynes XIV*. London: Macmillan.
- Khouri, Andrew, and Tracey Lien (2018). "Uber, Lyft and Other Gig Jobs May Face a Shakeup Under New California Work Rules." *Los Angeles Times*. May 02. Available at <http://www.latimes.com/business/la-fi-independent-contractors-20180502-story.html>.
- Lucas, Robert. 1978. "Unemployment Policy." *American Economic Review* 68: 353-7.
- Marshall, Thomas Humphrey. 1950. *Citizenship and Social Class*. London: Pluto Classics (1987 edition).
- Marx, Karl. 1867. *Capital*, Vol. 1. New York: Vintage (1977 edition).
- McPherson, Michael. 1983. "Review: Efficiency and Liberty in the Productive Enterprise: Recent Work in the Economics of Work Organization." *Philosophy and Public Affairs* 12: 354-68.
- Mill, John Stuart. 1848. *The Principles of Political Economy with Chapters on Socialism*. Jonathan Riley (ed.). Oxford: Oxford University Press (2008 edition).
- Nozick, Robert. 1974. *Anarchy, State, and Utopia*. New York: Basic Books.
- O'Donovan, Carolyn, and Jeremy Singer-Vine. 2016. "How Much Uber Drivers Actually Make Per Hour." *Buzzfeed*, June 22. Available at https://www.buzzfeed.com/carolineodonovan/internal-uber-driver-pay-numbers?utm_term=.gjQEgBOdY#.yWEzZW2Xr.
- Rawls, John. 1971. *A Theory of Justice*. Cambridge, MA: Harvard University Press.
- Rawls, John. 2001. *Justice as Fairness*. Erin Kelly (ed.). Cambridge, MA: Harvard University Press.
- Smith, Adam. 1776. *An Inquiry into the Nature and Causes of the Wealth of Nations*. Indianapolis, IN: LibertyClassics (1937 edition).
- Spencer, Herbert. 1884. *The Man Versus the State, with Six Essays on Government, Society and Freedom*. Eric Mack (ed.). Indianapolis, IN: LibertyClassics (1981 edition).
- Taylor, Robert. 2017. *Exit Left: Markets and Mobility in Republican Thought*. Oxford: Oxford University Press.
- Varian, Hal. 1976. "Keynesian Models of Unemployment." Working Paper No. 188, Department of Economics, Massachusetts Institute of Technology. Available at <https://archive.org/details/keynesianmodels000vari>.
- Walzer, Michael. 1978. "Town Meetings and Worker's Control: A Story for Socialists." *Dissent* 25 (1): 325-33.
- Wertheimer, Alan. 1990. *Coercion*. Princeton, NJ: Princeton University Press.
- Weil, David. 2014. *The Fissured Workplace: Why Work Became So Bad for So Many and What Can Be Done to Improve It*. Cambridge, MA: Harvard University Press.

B. WELFARE, RISK, AND POLICY

CHAPTER 19

COST-BENEFIT ANALYSIS AND SOCIAL WELFARE FUNCTIONS

MATTHEW D. ADLER

COST-BENEFIT analysis (CBA) and the social welfare function (SWF) framework are the two most prominent policy-analysis methodologies in economics. Both constitute systematic procedures for assessing governmental policies. CBA is widely used by governments (Renda 2011; Wiener 2013), and a large literature has arisen within applied economics to estimate the monetary valuations that are the input into CBA (Champ et al. 2003). The SWF framework, although not (yet) used by governments, has had a significant influence on academic economists, having been widely embraced in theoretical welfare economics. Moreover, it underpins scholarship on optimal tax theory; offers a normative foundation for inequality metrics; figures significantly in climate economics; and plays a role in other subfields of economics (Adler 2012: 57–88, 114–124; Tuomala 2016; Botzen and van den Bergh 2014; Cowell 2016).

The two approaches are similar in some important ways. Both are *consequentialist* and, more specifically, *welfarist*. By consequentialism I mean the ethical tradition that sees the goodness of outcomes as ethically fundamental.¹ The ethical status of a given act (such as a governmental policy choice) depends upon its associated outcome or, under uncertainty, its associated probability distribution across outcomes. Welfarism is the version of consequentialism that sees the moral goodness of outcomes as reducible to individual well-being (Adler 2012: 22–56).

CBA and the SWF framework both work within the tradition of welfarist consequentialism. Further, both adopt the preference account of individual well-being that is dominant in economics. On this account, whether a given person is better off in one outcome than a second depends upon whether she prefers the first to the second. Combining

¹ An outcome is a possible world or a cognitively tractable model of a possible world. Models are used by cognitively limited decision-makers in thinking about the goodness of worlds and policies.

welfarism with a preference account of well-being, we arrive at the axioms of Pareto Indifference and Strong Pareto (collectively, the “Pareto axioms” or “Pareto principles”):

Pareto Indifference: If each person is indifferent between x and y , then x and y are equally good.

Strong Pareto: If each person weakly prefers x to y , and at least one person strictly prefers x to y , then x is better than y .²

For a given set of outcomes, CBA and the SWF framework each generate a well-behaved, Paretian ranking of the outcomes: well-behaved in the sense that the ranking is complete and transitive, Paretian in the sense that the ranking satisfies the twin axioms of Pareto Indifference and Strong Pareto.

Notwithstanding these commonalities, CBA and the SWF approach differ in critical respects. The two methodologies often yield different rankings of a given set of outcomes and, thus, of a given set of policies. Such differences can arise because methodologies that converge in respecting the Pareto Indifference and Strong Pareto axioms can diverge in how they rank Pareto-noncomparable pairs of outcomes.³

This chapter provides a brief review of the two approaches. Their comparative merits are a topic of continuing debate in economics. This author falls in the camp that favors the SWF framework, and that perspective will be reflected in the review. Section 19.1 outlines the preference account of well-being and discusses utility functions, a tool that is used (in different ways) by both CBA and the SWF framework. Section 19.2 describes CBA, and Section 19.3 the SWF framework. Section 19.4 reviews the standard defenses of CBA, challenges these arguments, and in so doing defends the SWF framework as an improvement over CBA. This leaves open the issue of choosing a specific SWF, a topic addressed in Section 19.5. The *generalized utilitarian* family of SWFs emerges as especially plausible, including specifically the straight utilitarian SWF (whose formula is a simple summation of well-being numbers), and *prioritarian* SWFs that give greater weight to the well-being of the worse off.

CBA (as will be explained) means the unweighted sum of monetary equivalents. A different version of CBA adds up monetary equivalents multiplied by distributional weights (Adler 2016a; Boadway 2016). These weighting factors are used so that CBA approximates an SWF. Because distributionally weighted CBA is little used—the universal approach in government, and the dominant approach in the CBA literature, is to undertake unweighted CBA—it will not be discussed here.

A technical note: I assume throughout a “fixed population” framework. There is a set of N individuals, each of whom exists in all of the outcomes being compared. Variation in the identity of individuals or the size of the population is certainly a major problem for the theoretical elaboration of either approach, but does not bear upon the choice between the approaches, and will not be discussed in this summary presentation.

² To weakly prefer is to either strictly prefer or be indifferent. Throughout the chapter, “prefer” without a qualifier means strict preference.

³ Section 19.4.4 below discusses a special case in which CBA and the SWF framework converge.

19.1 WELL-BEING, PREFERENCES, AND UTILITY FUNCTIONS

What determines whether a given person (“Eve”) is better off, worse off, or equally well off in one outcome as compared to a second? This is philosophically contentious (Adler 2012: 155–84; Bykvist 2016; Fletcher 2016; Griffin 1986; Haybron 2016; Hurka 2016; Sumner 1996). The preference view, in its simplest version, says the following: Eve is at least as well off in x as compared to y iff—that is, if and only if—Eve weakly prefers x to y . This can be challenged from two directions. *Mental-state* accounts of well-being, such as hedonism, insist that Eve’s well-being supervenes on her mental states: if Eve has the very same mental states in x as in y , then she is equally well off in the two outcomes. Hedonism identifies the welfare-relevant mental states as hedonic states (pains and pleasures) and so says, specifically, that if Eve has the same hedonic states in x as in y , then she is equally well off in the two outcomes. Mental-state accounts are different from preference accounts, since an individual can prefer features of outcomes other than her own mental states—as famously illustrated by Robert Nozick’s “experience machine” (1974: 42–5).

Objective-good accounts of well-being challenge the preference view in a different way. Such accounts posit a list of goods at least some of which can be attained without being preferred. (Consider the good of knowledge: I can know lots of math despite having no yen for math.) On an objective-good account, Eve is at least as well off in x as compared to y iff Eve’s balance of goods is at least as good in x —whether or not she weakly prefers x .

Notwithstanding these challenges, a strong case can be made for the preference view. The view retains substantial philosophical support; was universally adopted by economists for much of the twentieth century; and is still dominant in economics. While some work in economics has moved either towards a mental-state view (as with happiness economics; see Graham 2016) or an objective-good view (as with work on “capabilities”; see Comim et al. 2008), this remains a minority approach. In particular, both CBA and the SWF framework are typically conjoined with the preference view.

The simple preference view needs refining. It actually is not plausible that Eve is at least as well off in x as y iff she weakly prefers x to y . Her preferences might be poorly informed or violate norms of rationality. Or, her preferences, albeit well-informed and rational, might be non-self-interested (as illustrated by Derek Parfit’s “stranger” example).⁴ The better version of the preference view says: Eve is at least as well off in x as y iff her well-informed, rational, and self-interested preferences are such as to weakly prefer x to y .

⁴ Parfit (1987: 494) imagines a case in which he meets a stranger, learns that the stranger has a disease, and out of sympathy for the stranger develops a preference that the disease be cured. The stranger is later cured, unbeknownst to Parfit. Although Parfit’s preference for a cure has been satisfied, Parfit is not better off.

Henceforth, I will assume these qualifiers: someone's preference just means her well-informed, rational, and self-interested preference. What this means is a complicated matter, beyond the scope of this chapter. Indeed, the CBA and SWF literatures have no consistent practice with respect to screening preferences for information, rationality, and self-interest. But these methodologies, to be ethically attractive, should incorporate these screens; and in what follows I will assume that this is done.

Let's now introduce some symbolism and turn to the topic of utility functions. I will use " R " to denote a preference. The arguments for preferences are bundles of individual attributes (income, health, longevity, leisure, social life, and so forth), with a whole bundle denoted as " a ," " b ," and so on. " $a R b$ " indicates that attribute bundle a is weakly preferred by preference R to attribute bundle b . An individual's preferences over outcomes, in turn, depend upon her attributes in the outcomes.⁵

A "lottery," denoted L , is a probability distribution over attribute bundles. $\pi_L(a)$ is the probability assigned to bundle a by lottery L . I will assume that R ranks lotteries as well as sure bundles. Indeed, if R ranks lotteries, it automatically ranks bundles too, because a bundle is simply a "degenerate" lottery—one assigning probability 1 to that bundle.

A "utility" function is a mathematical representation of a preference. So-called von-Neumann Morgenstern (vNM) utility functions are especially powerful (Gilboa 2009). They represent preferences over lotteries, not merely bundles, and do so in an "expectational" fashion. Whenever preference R weakly prefers lottery L to lottery L^* , the corresponding vNM utility function is such that its expected value with lottery L is at least as large as its expected value with lottery L^* . This chapter works solely with vNM utility functions. The symbol $u(\cdot)$ will be used to represent a vNM utility function, and $u^R(\cdot)$ the vNM utility function corresponding to preference R .

vNM utility functions are unique up to a positive affine transformation. If $u^R(\cdot)$ expectationally represents preference R with respect to lotteries over some set of attribute bundles, then $v^R(\cdot)$ also does so iff $v^R(\cdot)$ is a rescaling of $u^R(\cdot)$ by a change of origin and/or unit. That is, there exists a positive constant m and constant n such that $v^R(a) = mu^R(a) + n$ for all attribute bundles a .

19.2 CBA

On the theory of CBA, see generally Adler (2012: 88–114), Adler and Posner (2006), Boadway (2016), Boadway and Bruce (1984), Boardman et al. (2011), Freeman (2003), and Just et al. (2004). What follows describes CBA conjoined with a preference view of well-being; for CBA using instead a mental-state view, see Fujiwara and Dolan (2016).

⁵ Let $a_i(x)$ denote individual i 's bundle of attributes in outcome x . Then i weakly prefers x to y iff she weakly prefers $a_i(x)$ to $a_i(y)$.

The linchpin of CBA is what I will term the “monetary equivalent.” Consider two outcomes, x and y . Individual i ’s monetary equivalent for y , relative to x , is the amount of money—added to her holdings in x —that makes her indifferent between the two outcomes.

Assume individual i has preferences R . Let “ c ” denote individual i ’s consumption (monetary expenditure) in x , and “ d ” her non-consumption attributes. So her attribute bundle a in x is given by $a = (c, d)$. Let (c^*, d^*) be the individual’s bundle of consumption and non-consumption attributes in y . Then her monetary equivalent for y , relative to x , is the amount Δc such that: preferences R are indifferent between $(c + \Delta c, d)$ and (c^*, d^*) . In terms of utility functions: $u^R(c + \Delta c, d) = u^R(c^*, d^*)$. If the individual prefers y to x , Δc is positive; if she prefers x to y , Δc is negative. (A terminological note: the “monetary equivalent,” as here defined, is sometimes instead referred to in the CBA literature as a “willingness to pay” or “accept” amount, or as an “equivalent” or “compensating variation.”)

Consider now a set of outcomes $\{x, y, \dots\}$. Roughly speaking, the CBA rule is to sum monetary equivalents and to use that sum to rank the outcomes. But this rough rule can be fleshed out in various ways. Some ways of doing so run into technical difficulties; they can fail to yield an outcome ranking that is both well-behaved and Paretian.

These difficulties can be circumvented via the following rule. Arbitrarily choose a baseline outcome b . Let $ME_i(x)$ denote individual i ’s monetary equivalent for outcome x , relative to the baseline b . If we make certain assumptions about individual preferences for consumption, then a unique $ME_i(x)$ will exist for every individual i and outcome x . The key such assumption is non-satiation: individuals always prefer more consumption to less, *ceteris paribus*.⁶ Consider now the following rule (henceforth the “CBA rule”): outcome x at least as good as outcome y iff $\sum_{i=1}^N ME_i(x) \geq \sum_{i=1}^N ME_i(y)$. It is clear that this rule is well-behaved: the ranking of outcomes is complete and transitive. Further, it is Paretian:

Pareto Indifference: If each individual is indifferent between x and y , then $ME_i(x) = ME_i(y)$ for all i , and so the CBA rule assigns the same score to the two outcomes.

Strong Pareto: Note that, by nonsatiation: if i weakly prefers x to y , $ME_i(x) \geq ME_i(y)$; and if i strictly prefers x to y , $ME_i(x) > ME_i(y)$. Therefore, if everyone weakly prefers x to y , and some strictly prefer x , $\sum_{i=1}^N ME_i(x) > \sum_{i=1}^N ME_i(y)$.

Applying the CBA rule means estimating individuals’ monetary equivalents. This is far from straightforward; but a vast literature in applied economics has arisen to tackle the problem (Champ et al. 2003). Market prices, other types of behavioral evidence, and survey responses can all be brought to bear in the estimation task. Further, monetary equivalents are not literally ascertained individual-by-individual, but rather by placing

⁶ For every c, d and R : if $\Delta c^* > \Delta c$, $u^R(c + \Delta c^*, d) > u^R(c + \Delta c, d)$. We also have to assume that, whatever the individual’s bundle of attributes in any outcome x other than the baseline, there is some change to her baseline consumption that makes her indifferent between the baseline and x .

individuals in demographic groups assumed to have the same preferences and attributes. These implementation nuances are not essential to a basic understanding of CBA, or a comparison with the SWF approach, and will be ignored in what follows.

Note that the CBA rule does not require interpersonal comparisons. This feature of CBA is often invoked in its defense. The preference account of well-being is—in the first instance—an account of *intrapersonal* well-being comparisons. Recall that it says that Eve is at least as well off in x as y iff Eve weakly prefers x to y . So what is being analyzed is an intrapersonal fact about Eve's well-being—more precisely, an intrapersonal fact regarding Eve's well-being levels. Suppose now that interpersonal well-being comparisons are impossible; it is never the case that one person is at least as well off as a second, or that one person's well-being difference is as large as a second's. Even if this supposition is true, individuals' monetary equivalents will exist and the CBA rule will give a well-behaved, Paretian ranking of any set of outcomes as long as individuals have preferences meeting certain conditions (in particular, a non-satiation condition).

The CBA rule is not, yet, a policy assessment tool. It tells us how to rank outcomes; but governmental policymakers will often be uncertain about the outcomes that would result from the policy choices available to them. This can be captured by seeing policies as probability distributions over outcomes. The CBA policy tool will rank such distributions; and it will reduce to the CBA rule in the limiting case where, for each policy being evaluated, there is some outcome with probability 1 conditional on that policy.

Space limitations preclude an analysis of how the CBA outcome-ranking rule can be refined into a policy assessment tool. A discussion of the rule itself will suffice to give the reader a basic understanding of CBA, how it compares to the SWF framework, and how it might be justified.

19.3 SWFs

On the theory of SWFs, see generally Adler (2012, forthcoming), Blackorby et al. (2005), Boadway and Bruce (1984), Bossert and Weymark (2004), d'Aspremont and Gevers (2002), Mongin and d'Aspremont (1998), and Weymark (2016).

The SWF framework has two central components: an account of well-being and corresponding well-being measure, $w(\cdot)$, that converts outcomes into vectors (lists) of well-being numbers (one for each person in the population); and the "social welfare function" (SWF) itself, which is a rule for comparing these well-being vectors. A given outcome, x , is converted into the N -entry vector $(w_1(x), w_2(x), \dots, w_N(x))$. $w_1(x)$ is the well-being number of individual 1 in x , $w_2(x)$ the well-being number of individual 2 in x , and so forth. Outcomes are then ranked according to the ranking by the SWF of their corresponding well-being vectors.

The following symbolism will prove useful. Let $(w_1^*, \dots, w_N^*)$ and $(w_1, \dots, w_N)$ be two well-being vectors. $(w_1^*, \dots, w_N^*) \geq (w_1, \dots, w_N)$ indicates that the first vector is ranked at least as good as the second according to the SWF. Similarly, the symbol $>$ indicates "better than" according to the SWF, and $\sim$ indicates "equally good as" according to the SWF.

In short, the rule that the SWF framework uses to rank outcomes is this:

Outcome x is at least as good as outcome y iff $(w_1(x), w_2(x), \dots, w_N(x)) \succeq (w_1(y), w_2(y), \dots, w_N(y))$

This is a generic statement of the SWF framework, which is agnostic between different accounts of well-being. The specifics of marrying the framework with a preference account will be discussed below.

Note that the SWF framework as just presented is an outcome-ranking rule. This rule is not itself an adequate basis for policy assessment—given uncertainty about the nexus from policies to outcomes—just as the CBA rule is not. As with CBA, however we will generally ignore the nuances of refining the SWF framework into a methodology that allows us to rank policies, that is, probability distributions over outcomes.

What are the features of the SWF? At a minimum, the SWF is formally well-behaved, in the sense of yielding a complete and transitive ranking of any set of well-being vectors. Further, it is generally assumed that the SWF satisfies a *monotonicity* axiom:

Monotonicity: Let $(w_1, \dots, w_N)$ and $(w_1^*, \dots, w_N^*)$ be two well-being vectors. If $w_i^* \geq w_i$ for all i , and $w_j^* > w_j$ for at least one j , then $(w_1^*, \dots, w_N^*) \succ (w_1, \dots, w_N)$.⁷

Many different types of SWFs satisfy these basic conditions. We narrow down the space of possible SWFs by adding additional axioms—further requirements regarding the ranking of well-being vectors that we find ethically attractive. As we will see in Section 19.5, plausible axioms bring us to the *generalized utilitarian* family of SWFs, and yet more specifically to the utilitarian SWF and the prioritarian class of SWFs. The utilitarian SWF ranks vectors according to the simple sum of well-being, whereas a prioritarian SWF ranks vectors according to the sum of well-being numbers “plugged into” a concave transformation function.

The utilitarian SWF: $(w_1^*, \dots, w_N^*) \succeq (w_1, \dots, w_N)$ iff $\sum_{i=1}^N w_i^* \geq \sum_{i=1}^N w_i$

The prioritarian class of SWFs: $(w_1^*, \dots, w_N^*) \succeq (w_1, \dots, w_N)$ iff $\sum_{i=1}^N g(w_i^*) \geq \sum_{i=1}^N g(w_i)$, with $g(\cdot)$ some increasing and concave⁸ function⁹

What is the informational content of the well-being numbers assigned by the well-being measure $w(\cdot)$? The answer is that the numbers *represent all admissible well-being comparisons*—admissible as per the account of well-being that we have adopted and to which $w(\cdot)$ corresponds.

⁷ This monotonicity axiom is sometimes referred to as the Strong Pareto axiom. (Indeed, I have done so myself!) In this chapter, however, the term “Pareto” is reserved for axioms that specifically involve preferences, rather than well-being more generally.

⁸ “Increasing” means strictly increasing and “concave” means strictly concave.

⁹ More precisely, these are “continuous prioritarian” SWFs. In this chapter, I simply refer to them as “prioritarian.”

Table 19.1 Types of well-being comparisons

	Intrapersonal	Interpersonal
Levels	Admissible?	Admissible?
Differences	Admissible?	Admissible?

Four types of well-being comparisons are set forth in Table 19.1. A given account of well-being will allow for some or all of these four types of comparisons; it will recognize some or all of them as admissible (meaningful).

The most austere account admits only intrapersonal comparisons of well-being levels. On such an account, it is meaningful to compare Amy's well-being level in some outcome x to Amy's well-being level in outcome y —but not to compare Amy's well-being to that of Jose (a distinct individual); nor to compare Amy's difference in well-being between x and y to Amy's difference in well-being between z and zz ; nor to compare Amy's difference in well-being between x and y to Jim's difference in well-being between z and zz .

A different account might recognize both intra- and interpersonal comparisons of well-being levels, but refuse to admit difference comparisons. Or, the account might widen the austere account in a different way—allowing for intrapersonal comparisons of levels *and* differences, but eschewing all interpersonal comparisons. Finally, and most expansively, all four types of comparisons displayed in Table 19.1 might be seen as admissible.

Now we can better grasp the informational content of the well-being measure $w(\cdot)$. If intrapersonal comparisons of well-being levels are admissible, then: $w_i(x) \geq w_i(y)$ iff individual i in x is at least as well off as individual i in y . If intrapersonal comparisons of well-being differences are admissible, then: $w_i(x) - w_i(y) \geq w_i(z) - w_i(zz)$ iff the difference in i 's well-being between x and y is at least as large as the difference in i 's well-being between z and zz . If interpersonal comparisons of well-being levels are admissible, then: $w_i(x) \geq w_j(y)$ iff individual i in x is at least as well off as individual j in y . Finally, if interpersonal comparisons of well-being differences are admissible, then: $w_i(x) - w_j(y) \geq w_k(z) - w_l(zz)$ iff the difference in well-being between i in x and j in y is at least as large as the difference in well-being between k in z and l in zz .

Note that well-being measures are not *unique*. $w(\cdot)$ and $w^*(\cdot)$ may be two distinct measures, in the sense that the numbers assigned by $w(\cdot)$ are different from the numbers assigned by $w^*(\cdot)$, and yet contain the very same well-being information. For example, assume that all four boxes in Table 19.1 are admissible, according to the account of well-being we have adopted. Suppose that $w(\cdot)$ accurately represents the admissible well-being information. Consider now a different measure of well-being $w^*(\cdot)$, which is a common cardinal rescaling of $w(\cdot)$. That is, there exist two constants, $s > 0$ and t , such that $w_i^*(x) = sw_i(x) + t$ for all individuals i and outcomes x . Then it can be readily seen that this new measure, $w^*(\cdot)$, *also* represents the admissible well-being information—it is “informationally equivalent” to $w(\cdot)$.¹⁰

¹⁰ It is easy to see algebraically that $w^*(\cdot)$ contains exactly the same information about intra- and interpersonal comparisons of well-being levels and difference as $w(\cdot)$. To illustrate, consider the case of

It is very plausible to think that a given SWF—be it the utilitarian SWF, a prioritarian SWF, or any other—should be *invariant* to the substitution of informationally equivalent well-being measures. Assume that $w(\cdot)$ and $w^*(\cdot)$ are informationally equivalent. The SWF, coupled with measure $w(\cdot)$, produces a particular ranking of outcomes. Coupled with measure $w^*(\cdot)$, it produces a *different* ranking of outcomes. This would be highly problematic. There is no difference between the admissible well-being facts (concerning intra- and/or interpersonal level and/or difference comparisons) as represented by the first measure and those facts as represented by the second measure. And yet the SWF *varies* in how it ranks outcomes, depending on whether the SWF is paired with $w(\cdot)$ or instead with $w^*(\cdot)$. If so, the SWF is responding to meaningless numerical differences between $w(\cdot)$ and $w^*(\cdot)$ —differences that don't correspond to any genuine difference in the well-being facts.

A very important feature of the SWF framework is that it requires some degree of interpersonal comparability of well-being. We are now in a position to see why.

It is certainly possible to adopt a well-being account that crosses out the right column of Table 19.1—that admits only intrapersonal comparisons of well-being levels and differences—or yet more austere crosses out all cells except the top left. But denying interpersonal comparability produces serious difficulties for the SWF framework. Why? If only intrapersonal level and difference comparisons are admissible, and $w(\cdot)$ is a well-being measure, then any $w^*(\cdot)$ which is an individual-specific cardinal rescaling of $w(\cdot)$ is informationally equivalent to $w(\cdot)$.¹¹ *A fortiori*, if only intrapersonal level comparisons are admissible, and $w(\cdot)$ is a well-being measure, then any $w^*(\cdot)$ which is an individual-specific cardinal rescaling of $w(\cdot)$ is informationally equivalent to $w(\cdot)$. But, as Table 19.2 illustrates, neither the utilitarian SWF nor prioritarian SWFs are generally invariant to the replacement of $w(\cdot)$ with an individual-specific cardinal rescaling of $w(\cdot)$. The utilitarian SWF and prioritarian SWFs respond to well-being numbers as if these numbers do more than provide information about intrapersonal comparisons of well-being levels and differences. They mesh poorly with an account of well-being that crosses out the right column of Table 19.1.

Nor is this need for interpersonal comparability limited to the utilitarian and prioritarian SWFs. It turns out that *any* plausible SWF requires interpersonal comparability (of levels and/or differences). A key theorem of the SWF literature shows the following: if an SWF is invariant, in its ranking of outcomes, to the replacement of $w(\cdot)$ with an individual-specific cardinal rescaling of $w(\cdot)$, then either the SWF violates the Monotonicity axiom (a very compelling axiom) or the SWF must be a “serial dictatorship,” a flagrant violation of impartiality between individuals (Weymark 2016: 152).¹²

interpersonal comparisons of well-being differences. We need to show that $w_i(x) - w_j(y) \geq w_k(z) - w_l(zz)$ iff $w_i^*(x) - w_j^*(y) \geq w_k^*(z) - w_l^*(zz)$ for every i, j, k, l , and every x, y, z, zz . Because $w_i^*(x) = s_i w_i(x) + t_i$ for every x, i , with $s_i > 0$, this is clearly true.

¹¹ A well-being measure $w^*(\cdot)$ is an individual-specific cardinal recalling of $w(\cdot)$ iff: there exist two constants $s_i > 0$ and t_i for each individual i , such that: $w_i^*(x) = s_i w_i(x) + t_i$ for all individuals i and outcomes x .

¹² A serial dictatorship orders the N individuals in degrees of dictatorial power. Two vectors are ranked according to the well-being levels of the most powerful individual; if these are equal, according to the second most powerful; and so forth.

Table 19.2 Invariance to rescalings of the well-being measure

Original well-being numbers				Individual-specific cardinal rescaling					
Outcome				Outcome			Scaling Factors		
	x	y	z		x	y	z	s_i	t_i
Abel	9	4	21	Abel	9	4	21	1	0
Bob	25	25	21	Bob	230	230	190	10	-20
Cindy	49	16	21	Cindy	7.9	4.6	5.1	0.1	3
Diana	1	36	21	Diana	400	14,400	8400	400	0
				Order					Order
Sum of w_i	84	81	84	$x \sim z > y$	646.9	14,638.6	8616.1	$y > z > x$	
Sum of $\sqrt{w_i}$	16	17	18.3	$z > y > x$	41	139.3	112.3	$y > z > x$	

Explanation: The left-hand side of the table displays well-being numbers as assigned by some measure $w(\cdot)$, while the right-hand side shows an individual-specific cardinal rescaling of these numbers, with the scaling factors in the rightmost two columns. The bottom rows of the table illustrate the ranking of the outcomes according to the utilitarian SWF and the sum-of-square-root SWF, an exemplary prioritarian SWF. The order of the outcomes x, y, z —according to each SWF combined with $w(\cdot)$ or the rescaled measure—is indicated under the heading “order.” The symbol “ $\sim$ ” indicates that the outcomes are equally good, and “ $>$ ” that the first outcome is better than the second. As the table shows, each of the SWFs changes its order of the outcomes as we shift from the well-being measure $w(\cdot)$, on the left side of the table, to the rescaled measure on the right.

So the proponent of the SWF framework faces a challenge: developing an account of well-being that allow for interpersonal comparability, and constructing a corresponding well-being measure $w(\cdot)$. This is a generic challenge, which must be surmounted regardless of which type of well-being account is adopted. I’ll now outline how the challenge can be surmounted on a preference view (based upon Adler 2016b).

The core of the preference account is an analysis of intrapersonal comparisons of well-being levels (the top left cell of Table 19.1) in terms of individual preferences: Sondra is at least as well off in x as y iff Sondra weakly prefers x to y . How are we to expand the account, so as to preserve this core, yet also make sense of interpersonal comparisons?

Consider first the case in which individuals have *common* preferences over attribute bundles and bundle lotteries. Denote this common ranking as R : Amy ranks bundles and lotteries according to R , and so does Bob, Cindy, Daniel, and every other individual in the population. Let $u^R(\cdot)$ be a vNM utility function representing preference ranking R . We allow for all four types of comparisons in Table 19.1, which are analyzed as follows.

- (1) Individual i in x is at least as well off as individual i in y iff R is such as to weakly prefer individual i ’s bundle in x to his bundle in y .
- (2) Individual i in x is at least as well off as individual j in y iff R is such as to weakly prefer individual i ’s bundle in x to individual j ’s bundle in y .

- (3) Intrapersonal well-being differences are analyzed in terms of differences in vNM utility. That is, the difference in individual i 's well-being between x and y is at least as large as the difference in her well-being between z and zz iff $u^R(a_i(x)) - u^R(a_i(y)) \geq u^R(a_i(z)) - u^R(a_i(zz))$.
- (4) Interpersonal well-being differences are similarly analyzed in terms of differences in vNM utility. The difference in well-being between individual i in x and individual j in y is at least as large as the difference in well-being between individual k in z and individual l in zz iff $u^R(a_i(x)) - u^R(a_j(y)) \geq u^R(a_k(z)) - u^R(a_l(zz))$.

A well-being measure $w(\cdot)$ representing all four types of comparisons is constructed by defining $w(\cdot)$ in terms of the vNM utility function $u^R(\cdot)$, using the following formula: $w_i(x) = u^R(a_i(x))$.¹³ In other words, the well-being number of individual i in outcome x is just the utility number of her attributes in x , as assigned by the vNM utility function $u^R(\cdot)$ capturing the common preference structure R . The well-being number $w(\cdot)$ is not unique, but unique up to a common cardinal rescaling. Recall that vNM utility functions are unique up to positive affine transformations: if $u^R(\cdot)$ represents preference R , then so does $v^R(\cdot)$, if $v^R(\cdot) = mu^R(a) + n$. The upshot is that a well-being measure defined using the formula immediately above is unique up to a common cardinal rescaling.

For short, let's call the account of well-being and its measurement described in the previous two paragraphs the "vNM approach." Its analysis of interpersonal comparisons of well-being levels, prong (2) above, should be relatively uncontroversial. On a preference view of well-being, there are two ingredients in an individual's welfare: what she wants (her preferences) and what she has (her attributes). If two individuals have the same wants, then a comparison of their well-being levels should be nothing other than a comparison of what they have (their attributes), as viewed through the lens of the common wants (the common R).

The supposition that we can make difference comparisons, intra- or interpersonal, is intuitively plausible. Jim has an intense migraine and a mildly scratchy throat. I can give Jim a pill to stop the migraine or a lozenge for the throat. It is natural to say that the pill makes a bigger difference to Jim's welfare than the lozenge, and pedantic to insist that we cannot compare the two well-being changes. It is also pedantic to allow that the well-being difference from giving Jim a pill to relieve his migraine is larger than the well-being difference from giving him a lozenge to relieve his scratchy throat, but to insist that the well-being difference from giving Jenna a pill to relieve her migraine is incomparable with the well-being difference from giving Jim the lozenge.

A more contestable feature of prongs (3) and (4) above is the specific analysis of well-being differences in terms of differences in vNM utility. Space constraints preclude a defense here of this feature of the vNM approach. I believe it to be quite plausible.

¹³ It is clear that $u^R(\cdot)$ represents intra- and interpersonal comparisons of well-being levels, as defined in the previous paragraph. Note that $a R b$ iff $u^R(a) \geq u^R(b)$. As for intra- and interpersonal comparisons of well-being differences: these are analyzed in terms of numerical differences in vNM utility values, and so the representation is immediate.

Let's return to the formula by which the vNM approach constructs a well-being measure: $w_i(x) = u^R(a_i(x))$. It is not too hard to see that this well-being measure, if used in conjunction with an SWF satisfying the Monotonicity axiom, will produce a ranking of outcomes satisfying both Pareto Indifference and Strong Pareto.¹⁴ The SWF framework does not *automatically* satisfy the Pareto principles—we need to use an appropriate well-being measure, namely one that respects preferences, and to have a monotonic SWF—but if we do use an appropriate well-being measure, in particular the vNM measure, and do use a monotonic SWF, the Pareto principles are satisfied.

In presenting the vNM approach, I have focused on the simple case of common preferences over attribute bundles and lotteries. The approach smoothly generalizes to the case of heterogeneous preferences. Among the population of individuals, some have preference R , some have preference R^* , and so forth. For each preference, we choose a corresponding utility function $u^R(\cdot)$ plus scaling factors. The collection of utility functions and scaling factors is used to construct a well-being measure $w(\cdot)$. The scaling factors are needed to make comparisons of well-being levels and differences between individuals with different preferences. Various techniques can be used to identify the scaling factors: for example, a “high”/“low” scaling, whereby individuals regardless of their preferences are assumed to have the same well-being level at a “high” (good) attribute bundle and “low” (bad) bundle. The well-being measure $w(\cdot)$, conjoined with a monotonic SWF, will—as in the simpler case of common preferences—yield an outcome ranking that satisfies the Pareto principles. (See Adler 2016b for details.)

19.4 DEFENSES OF CBA

CBA is widely used by governments to assess policies. Is this justified? Here I review standard defenses of CBA. I do so not in the abstract, but rather comparatively. An alternative policy-analytic framework—the SWF approach—is well developed in the academic literature. With this competitor “waiting in the wings,” so to speak, why rely on CBA?

When comparing CBA to the SWF approach, I assume (unless otherwise noted) that the SWF is monotonic and coupled with a preference-respecting measure of well-being,¹⁵ specifically one constructed using the vNM approach.

While CBA and the SWF framework both satisfy the Pareto principles, they still differ in significant ways. Consider two outcomes x and y that are “Pareto noncomparable”: at least one individual strictly prefers x to y , and at least one individual strictly prefers y to x .

¹⁴ *Pareto Indifference.* Assume that each person is indifferent between x and y . Then $u^R(a_i(x)) = u^R(a_i(y))$ for all i . Then, by the formula immediately above, $w_i(x) = w_i(y)$ for all i . Thus the well-being vectors for the two outcomes are the same, and any SWF will rank the outcomes the same.

Strong Pareto. Assume that each person weakly prefers x to y and some strictly prefer it. Then $u^R(a_i(x)) \geq u^R(a_i(y))$ for all i and $u^R(a_j(x)) > u^R(a_j(y))$ for at least one j . Thus $w_i(x) \geq w_i(y)$ for all i and $w_j(x) > w_j(y)$ for at least one j . If the SWF satisfies Monotonicity, it ranks x above y .

¹⁵ That is, $w_i(x) \geq w_i(y)$ iff i weakly prefers x to y .

The Pareto principles do not constrain the x/y ranking; it is consistent with those principles to rank x over y , y over x , or the two equally good. CBA and the SWF framework can, and often do, diverge in their ranking of Pareto-noncomparable pairs of outcomes.

Here is a simple example of such divergence. Assume that individuals have common preferences R and—as standardly assumed—are risk averse with respect to consumption lotteries.¹⁶ This means that a well-being measure $w(\cdot)$ constructed using the vNM method will exhibit diminishing marginal utility of consumption. If Kayla and Lanny are otherwise identical, but Kayla's level of consumption is lower than Lanny's, augmenting Kayla's consumption by Δc will produce a bigger increase in well-being than augmenting Lanny's by the same amount.

It therefore follows that both the utilitarian SWF and prioritarian SWFs favor an equalization of consumption, *ceteris paribus*. Let y be an outcome identical to x , except that Δc has been transferred from Lanny to Kayla.¹⁷ Note that the outcomes are Pareto noncomparable: Kayla prefers y to x , while Lanny prefers x to y . The utilitarian SWF and prioritarian SWFs will prefer y to x , while CBA may be indifferent. In particular, assume that x is the baseline outcome used in the calculation of monetary equivalents. Then Kayla's monetary equivalent for y is Δc , while Lanny's is $-\Delta c$, and the sum is zero.¹⁸

19.4.1 Kaldor-Hicks Efficiency and the Impossibility of Interpersonal Comparisons

The construct of Kaldor-Hicks (KH) efficiency was developed during the 1930s and 1940s (Hicks 1939; Kaldor 1939; Hotelling 1938), a time when leading economists were skeptical of interpersonal comparisons and were seeking to develop assessment tools that did not require them. There are various, subtly different definitions of KH efficiency, but they all appeal to the *potential* of converting some outcome into a Pareto-superior outcome relative to another.¹⁹ The following definition will suffice for our purposes: outcome x is KH efficient relative to outcome y if there is some outcome x' which can be reached from x by a lump-sum transfer of resources and is Pareto superior to y . Note that KH efficiency indeed eschews interpersonal comparisons. Even if such comparisons are impossible, we can apply the KH criterion as long as we know each person's ranking of outcomes.

One defense of CBA appeals to KH efficiency: If x is preferred to y by CBA, then x is KH efficient relative to y , and so better than y . Or so the defense goes (Adler 2012: 98–104).

¹⁶ By consumption lotteries, I mean lotteries over attribute bundles in which the consumption attribute varies but other attributes are held fixed.

¹⁷ I assume that Δc is less than the consumption gap between the two individuals.

¹⁸ Complexities arise if Kayla or Lanny have different non-consumption attributes in the baseline than in x and y . In that case, CBA need not be indifferent to the transfer.

¹⁹ "Pareto superior" just signals the applicability of the Strong Pareto axiom: one outcome is Pareto superior to a second if everyone weakly prefers the first and at least some strictly prefer it.

One difficulty with this defense of CBA is technical: CBA and KH efficiency are not logically equivalent. Consider the CBA rule described earlier: outcome x at least as good as outcome y iff $\sum_{i=1}^N ME_i(x) \geq \sum_{i=1}^N ME_i(y)$, with monetary equivalents defined as compensating changes to consumption in baseline b . It is not too hard to see that x may be preferred to y by the CBA rule and yet not KH efficient relative to y . Why? The definition of KH efficiency depends upon the potential of lump-sum redistribution in x to render it Pareto superior to y . But the CBA rule calculates monetary equivalents in b , not in x . The effect of consumption and consumption transfers in b on individuals' utilities may be different than in x .

To be sure, the CBA rule as just defined is merely one specification of the general idea of CBA (the sum of monetary equivalents). Similarly, the specific definition of KH efficiency offered several paragraphs above is merely *one* version of a more generic notion (potential Pareto superiority). But there seems to be no specific definition of CBA, and no specific definition of KH efficiency, such that: x is better than y according to CBA iff x is KH efficient relative to y . Boadway and Bruce, reviewing this issue, conclude: "The use of the unweighted sum of household compensating or equivalent variations [i.e., monetary equivalents] as a necessary and sufficient indicator of potential Pareto improvements is rife with difficulties" (1984: 271).

This technical divergence between CBA and KH efficiency is a significant problem for the KH defense of CBA. But let's assume that the problem can be circumvented, by seeing CBA as a reliable if not perfect indicator of KH efficiency. Even so, the KH defense of CBA has serious flaws.

First, its skepticism about interpersonal comparability is unwarranted. In an influential book that was part of the backdrop for the development of the KH criterion, the economist Lionel Robbins (1932) famously argued that interpersonal comparisons are not required for economics as a social science. This is true—but the KH criterion, CBA, and the SWF framework all transcend economics as a social science. These are *normative* tools, designed to give *guidance* in assessing policies—designed to tell policymakers what they "ought" to do. The irrelevance of interpersonal comparisons for purposes of economic science does not imply their irrelevance for purposes of policy assessment.

A closely related argument developed by Robbins is that interpersonal comparisons are not reducible to empirical facts. Simply knowing the facts (including facts about individuals' preferences) will not suffice to justify the conclusion that one person is better off than a second. Making interpersonal comparisons requires normative judgments. But—just because policy assessment *is* normative—one should hardly expect that the specification and application of policy tools will be devoid of normative judgments. For example: the KH criterion, CBA, and the SWF framework are all designed to conform to the principles of Pareto Indifference and Strong Pareto. Endorsing these principles involves various normative judgments: that well-being consists in preference satisfaction, that two outcomes equally good for well-being are equally ethically good (welfarism), and that increasing one person's well-being without decreasing anyone else's always increases ethical goodness.

Interpersonal comparability might be attacked in a different way. It might be claimed that, as a normative matter, such comparisons are incoherent or problematic. But this, too, is unpersuasive. Interpersonal comparisons of both well-being levels and differences are a routine part of lay discourse. Providing a precise account of such comparisons poses some challenges—but providing a precise account of *much* of lay moral discourse is a challenge! The vNM approach to interpersonal comparability is one route to surmounting the challenges.

Second, regardless of one's view about interpersonal comparability, the KH criterion is a problematic basis for ranking outcomes. Consider any case in which x and y are Pareto-noncomparable (some prefer x to y , others y to x), and x is KH efficient relative to y . Why think in this case that x is better than y ? By hypothesis, x can by lump-sum redistribution of resources be converted into x' , which is Pareto superior to y . But this shows that x' is better than y , not that x itself is. It is an odd piece of normative reasoning indeed to think that the transformability of one item into a second allows one to ascribe the ethical properties of the second to the first—in this case, to leap from the premises that (1) x can be transformed into x' via lump-sum redistribution and (2) x' is better than y , to the unwarranted conclusion that (3) x is better than y . If one is skeptical about interpersonal comparability, one's verdict about this case should be that x and y are ethically noncomparable. If one is not skeptical, an SWF can be used to rank x versus y —not the KH criterion.

19.4.2 Everyone Benefits from CBA in the Long Run

A different defense of CBA claims that it is universally beneficial in the long run (Adler 2012: 566–7). This line of argument defends CBA at the level of institutional structure. It does not contend that the CBA rule itself is an appropriate criterion of the ethical goodness of outcomes—that y is better than x whenever preferred to x by the CBA rule. Rather, the contention is that an institutional structure in which CBA is used to select policies is better for everyone than an alternative structure. Some individuals will lose from any particular policy choice, but in each such case there is a surplus of monetized gains over monetized losses—and these are “spread around” in the long run, to everyone's advantage.

While the KH-efficiency defense of CBA appeals to potential Pareto superiority, this defense appeals to *actual* Pareto superiority. The cumulative effects of CBA, over time, are such as to make the institution of CBA *actually* Pareto superior (not merely potentially so) as compared to the alternative.

This contention, in turn, can be advanced in two different ways. Let I^{CBA} be an institutional structure in which CBA is used by policymakers to select policies, and I^O a structure without CBA. One version of the long-run argument for CBA is this: The outcome of I^{CBA} is Pareto superior to the outcome of I^O . A different version relies on expected benefit: Each person's expected well-being with I^{CBA} is greater than her expected well-being with I^O .

On closer inspection, neither version of the long-run argument for CBA is persuasive. There are *many* different institutional alternatives to the use of CBA. We need to compare CBA as an institution (I^{CBA}) to *all* the plausible alternatives, not merely a single I^0 . Let I^0 be a “night watchman” state in which government does very little policymaking (merely enforcing fundamental rights of person, property, and contract). There is some plausibility in the thought that I^{CBA} is better for everyone than I^0 thus defined. But a different type of institutional alternative, clearly, is a non-night-watchman state in which some policy assessment tool *other than* CBA is used to assess policies. For example, let I^U be an institutional structure in which the utilitarian SWF is used to assess policies; I^{Prior} in which some prioritarian SWF is used to assess policies; I^{CE} in which the non-CBA technique of cost-effectiveness analysis is used to evaluate policies; and so on. There’s no reason to think that I^{CBA} is better for each person than every one of these alternatives.

19.4.3 CBA Is a Rough Proxy for Overall Well-Being

In our 2006 book, Eric Posner and I offered a novel defense of CBA. We endorsed rather than rejecting interpersonal comparability, and then argued that CBA should be seen as a rough proxy for overall well-being. Given the diminishing marginal utility of consumption, CBA is not equivalent to overall well-being, but under certain conditions, CBA and overall well-being will converge. Most simply, assume that everyone has the same marginal utility of consumption in the baseline outcome b . Then the ranking of outcomes by the CBA rule will be the same as their ranking in light of overall well-being, if monetary equivalents are sufficiently small.

More generally, in an x/y comparison, with some having larger monetary equivalents for x than for y —they benefit from x —and others having smaller monetary equivalents, the CBA rule will be a rough if not perfect proxy for overall being if the distribution of baseline marginal utility among the first group is roughly the same as the distribution of baseline marginal utility among the second group.

The ethical standard of “overall well-being,” to which Posner and I appealed, is what the utilitarian SWF makes precise: the sum total of interpersonally comparable individual well-being numbers. The “rough proxy” defense of CBA will, therefore, have little persuasive force for those inclined to favor a non-utilitarian SWF. For example, prioritarians do not care about overall well-being, but instead about priority-adjusted well-being (as formally expressed by the sum of concavely transformed well-being). Even if CBA were a perfect proxy for the utilitarian SWF, prioritarians would not favor CBA.

But what about utilitarians? A utilitarian should be persuaded by the rough proxy defense of CBA *if* she thinks that CBA is the best methodology available to policymakers to operationalize the standard of overall well-being—that nothing better is on hand. But I have now come to believe that there is a better methodology on hand, namely the SWF

framework. The policymaker can directly apply the utilitarian SWF, or use CBA with utilitarian distributional weights. It is not clear why—given these alternatives—CBA without distributional weights is justified.

19.4.4 CBA as an SWF

Why not see CBA itself as an SWF? Consider again the CBA rule: x at least as good as y iff $\sum_{i=1}^N ME_i(x) \geq \sum_{i=1}^N ME_i(y)$. Equivalently, the CBA rule is: x at least as good as y iff $\sum_{i=1}^N (ME_i(x) - ME_i(y)) \geq 0$. Consider now the utilitarian SWF: x at least as good as y iff $\sum_{i=1}^N w_i(x) \geq \sum_{i=1}^N w_i(y)$ or, equivalently, $\sum_{i=1}^N (w_i(x) - w_i(y)) \geq 0$.

Let $w^+(\cdot)$ be a measure of well-being that has the following property: for every x, y , and every individual i , $w_i^+(x) - w_i^+(y) = ME_i(x) - ME_i(y)$. Then CBA is equivalent to a utilitarian SWF using $w^+(\cdot)$ as its well-being measure. $w^+(\cdot)$ is a well-being measure with the special property that differences in each individual's $w^+(\cdot)$ values, between any two outcomes, are equal to differences in her monetary equivalents.

A well-being measure of this type can't be constructed using the vNM approach. If individuals are risk averse with respect to consumption lotteries (as typically assumed), a vNM well-being measure exhibits diminishing marginal utility of consumption. By contrast, $w^+(\cdot)$ does not exhibit diminishing marginal utility of consumption. However, there is a different method for constructing well-being measures, the so-called “equivalent income” approach, that can be used to construct $w^+(\cdot)$.²⁰

It's important to be clear about how the defense of CBA now on the table is different from the “rough proxy” defense discussed immediately above. That defense presupposes that well-being *does* exhibit diminishing marginal utility of consumption but observes that, under appropriate conditions, CBA is at least a rough proxy for the utilitarian SWF. The defense now on the table says: given the measure $w^+(\cdot)$, CBA is a *perfect* proxy for the utilitarian SWF.

The difficulties with this defense of CBA are twofold. First, the equivalent-income method for constructing a well-being measure can be challenged on various grounds. One of these, which is immediately to the point, is that the failure to exhibit diminishing

²⁰ An individual's “equivalent income” is her actual income adjusted for differences between her actual non-income attributes and a reference level of non-income attributes (Fleurbaey 2016). In this chapter, I am not distinguishing between income and consumption, and so I can offer a definition in terms of consumption. Let d^{ref} be the reference level of non-consumption attributes. Then an individual's equivalent consumption for the bundle (c, d) , denoted c^{equiv} , is such that she is indifferent between $(c^{\text{equiv}}, d^{\text{ref}})$ and (c, d) . Let $w_i^+(x)$ just be set equal to i 's equivalent consumption for her bundle of consumption and non-consumption attributes in x .

To define monetary equivalents, pick the baseline so that each individual has non-consumption attributes at d^{ref} , and at some arbitrary consumption level, c_i for individual i , which can vary among individuals. By the definition of monetary equivalent, individual i 's equivalent consumption for any outcome x is just $c_i + ME_i(x)$. Thus $w_i^+(x) - w_i^+(y) = ME_i(x) - ME_i(y)$ for all i, x, y .

marginal utility is counterintuitive. Intuitively, the benefit to a millionaire of increasing her income by \$1000 is smaller than the benefit to someone with a middle-class income, in turn smaller than the benefit to someone who is impoverished.

Second, a $w^+(\cdot)$ measure of well-being together with the utilitarian SWF implies that the SWF is insensitive to the distribution of consumption. Specifically, there's no ethical benefit in equalizing the baseline distribution of consumption. This unpleasant result can be avoided by combining $w^+(\cdot)$ with a non-utilitarian SWF.²¹ In particular, a prioritarian SWF together with $w^+(\cdot)$ will favor the equalization of consumption. But this combination is not equivalent to CBA.

19.4.5 The Tax System

An important argument in favor of governmental use of CBA has been developed most rigorously by Louis Kaplow (1996, 2004, 2008); for a more detailed discussion of Kaplow's view, see Adler (2012: 560–6, 2017). The argument is *not* skeptical about interpersonal comparability; rather, it is designed to *appeal* to those who endorse the SWF framework. Nor does it rely upon a special well-being measure (the $w^+(\cdot)$ measure), or see CBA as merely a rough proxy for an SWF. Instead, the Kaplow analysis argues for corralling the SWF framework into a specific institutional role. SWFs can and should be used in the design of the income tax system. But CBA, rather than an SWF, should be employed as the governmental policy framework for assessing all non-tax policies—regulations, building infrastructure, public goods provision, and all other non-tax policies.

The argument runs as follows. Assume policy P^* is preferred to policy P by CBA, while the policymaker's favored SWF (whatever it may be) prefers P to P^* . Then, under appropriate conditions, there will exist a change to the income tax system (let's denote this as ΔT), such that everyone can expect to be better off with P^* together with ΔT as compared to P .

What are “appropriate conditions”? They include the following: (1) the characteristics that determine whether individuals can expect to be better or worse off with P^* versus P must be identifiable to taxing authorities, or at least be correlated with identifiable characteristics; (2) individuals' preferences as between leisure and other attributes must have an appropriate form, so that the change ΔT does not induce a change in labor supply; (3) the administrative cost of changing the tax code and administering the new provisions must not be too high.

It is empirically contingent whether such conditions obtain. Further, even if they *do* obtain, that does not guarantee, of course, that the enactment of P^* will actually be followed by ΔT . That depends upon the political economy of the income tax system.

For these reasons, I do not believe that the Kaplow argument succeeds in demonstrating that CBA should be used for non-tax policies. Rather, what it shows is something

²¹ Indeed, the proponents of the equivalent-income approach to measuring well-being typically do not favor a utilitarian SWF.

weaker: policymakers should use SWFs to assess non-tax policies but should be sensitive to the possibility of combining these policies with beneficial alterations to the tax system. Assume that the policymaker favors some SWF (whatever it may be) and that the SWF prefers P to P^* ; however, there appears to be a change to the income system ΔT such that P^* plus ΔT is preferred by the SWF to P . If there indeed exists such a ΔT , and if the policymaker believes that the enactment of P^* will in fact be followed by the enactment of ΔT , then the policymaker should choose P^* . And this is exactly what the SWF framework tells the policymaker. The SWF framework tells her to choose P^* *under these conditions*, and otherwise not. By contrast, if the policymaker chooses P^* over P whenever CBA recommends, she will choose P^* even if the result of enacting P^* will *not* be a hybrid policy P^* plus ΔT that the SWF prefers to P .

19.5 WHAT TYPE OF SWF?

The SWF framework, to repeat, combines a measure $w(\cdot)$ of well-being that converts outcomes into vectors of well-being numbers; and a rule (the SWF) for ranking these vectors. We have discussed the construction of $w(\cdot)$. As for the SWF, we have assumed that it is formally well-behaved and satisfies a Monotonicity axiom. Monotonicity, together with a preference-respecting $w(\cdot)$, yields a ranking of outcomes that satisfies the Pareto principles.

A further very plausible axiom is Anonymity.

Anonymity: Let the well-being vector $(w_1, w_2, \dots, w_N)$ be a permutation (rearrangement) of the well-being vector $(w_1^*, w_2^*, \dots, w_N^*)$. Then the two vectors are equally good.

Anonymity is a formal expression of ethical impartiality. It should not matter that Amy is at level 20, Bob at level 3, and Carla at level 15, as opposed to Bob being at level 20 and Amy at 15 and Carla at 3.

There is a vast space of SWFs that satisfy Monotonicity and Anonymity. How should we choose within this space? We can do so by continuing to work with the axiomatic method: positing additional axioms that seem reasonable, either on fundamental normative grounds or as a pragmatic matter, and seeing which SWFs they imply (Adler 2012: ch.5, forthcoming).

Two axioms that have large pragmatic advantages are *Person Separability* and *Continuity*. I will provide an informal statement of each. Let's say that an individual is "unaffected" between a pair of vectors if her well-being level is the same in each.

Person Separability. Let $(w_1, \dots, w_N)$ and $(v_1, \dots, v_N)$ be one pair of well-being vectors, and $(w_1^*, \dots, w_N^*)$ and $(v_1^*, \dots, v_N^*)$ a second. The individuals unaffected between the first pair are the same as those unaffected between the second; and the only difference between the pairs is that each such unaffected person may be at a different well-being

level in the second pair than in the first. Then the SWF's ranking of the first pair is the same as its ranking of the second.

Continuity. If one vector is better than or worse than a second, this remains true for a sufficiently small region around the first.

Person Separability, generalized to cover policies (probability distributions over outcomes), means that the decision-maker does not need to gather information about unaffected persons in determining which policy is better. She can train her attention on a narrower question: how will the policies change the well-being of those individuals who stand to gain or lose? Continuity, too, is very useful. It means that the SWF can be represented by a real valued function, indeed a continuous function. In other words, there is some continuous real-valued $e(\cdot)$ such that $(w_1, \dots, w_N) \succcurlyeq (w_1^*, \dots, w_N^*)$ iff $e(w_1, \dots, w_N) \geq e(w_1^*, \dots, w_N^*)$.

Putting together Monotonicity, Anonymity, Person Separability and Continuity, we arrive at the *generalized utilitarian* family of SWFs. If an SWF satisfies these axioms, then it is represented by a mathematical function that is not merely continuous, but additive. $(w_1, \dots, w_N) \succcurlyeq (w_1^*, \dots, w_N^*)$ iff $\sum_{i=1}^N h(w_i) \geq \sum_{i=1}^N h(w_i^*)$, with $h(\cdot)$ strictly increasing and continuous.

As already mentioned, two types of generalized utilitarian SWFs are widely used in the SWF literature: the utilitarian SWF (using the formula $\sum_{i=1}^N w_i$) and prioritarian SWFs (using the formula $\sum_{i=1}^N g(w_i)$, with the $g(\cdot)$ function increasing and concave). Although other types of generalized utilitarian SWFs also exist, these seem less attractive normatively and, indeed, have been pretty much ignored.

The utilitarian SWF corresponds, of course, to the philosophical tradition of utilitarianism (Eggleston and Miller 2014), dating back centuries to Jeremy Bentham. Utilitarianism is now understood, by philosophers, to be an ethical theory with the following features. The theory is *consequentialist* (as opposed to deontological ethics or other varieties of ethical non-consequentialism). More specifically, it is *welfarist* (reducing the goodness of outcomes to individual well-being). Yet more precisely, utilitarianism is *simply additive*, that is, ranking outcomes according to the simple sum total of well-being. Bentham tightened the screw several more turns—adopting a mental-state, indeed hedonic view of well-being, as pleasure and the absence of pain. Since then, the tradition has broadened to include the full spectrum of well-being accounts. There are preference utilitarians and objective-good utilitarians as well as mental-state utilitarians. These possibilities are mirrored by the SWF format. As already explained, the format is generic in the sense that the $w(\cdot)$ numbers plugged into the utilitarian-SWF formula, $\sum_{i=1}^N w_i$, can represent well-being according to *any* account (as long as it permits numerical measurement and recognizes inter- as well as intrapersonal comparisons).

Prioritarian SWFs correspond to *prioritarianism*, a theory introduced to philosophy by Derek Parfit (2000) and Larry Temkin (1993: ch. 9) in the 1990s (see Holtug 2017). Prioritarianism is a type of welfare-consequentialism which is neither egalitarian nor utilitarian. Welfare egalitarians give intrinsic moral weight to equality, in the following sense: Whether x is morally better than y depends upon how each individual fares,

relative to others, in x , and on how each individual fares, relative to others, in y . Prioritarians do not give intrinsic moral weight to the equality of well-being. As Parfit explains, prioritarians “do not think it in itself bad, or unjust, that some people are worse off than others” (2000: 104). They thereby avoid the so-called “levelling down” objection: that an equalization of well-being that occurs by way of a *reduction* in the well-being of better-off individuals, to bring them closer to the levels of the worse-off, with no increase in anyone’s welfare, must be counted by the welfare egalitarian as an ethical improvement—if not an all-things-considered improvement, then at least a *pro tanto* improvement.

Prioritarians eschew an unwelcome feature of utilitarianism, namely that utilitarians always prefer to give a larger benefit to a better-off person than to give a smaller benefit to a worse-off person, regardless of how close in magnitude the benefits and how disparate the well-being levels of the two individuals. As Parfit writes,

For Utilitarians, the moral importance of each benefit depends only on how great this benefit would be. For *Prioritarians*, it also depends on how well off the person is to whom the benefit comes. We should not give equal weight to equal benefits, whoever receives them. Benefits to the worse off should be given more weight. (2000: 101)

The prioritarian-SWF formula, $\sum_{i=1}^N g(w_i)$, does indeed give priority to the worse off—because the formula is not straight additive, but instead sums an increasing and *concave* transformation of individual well-being. Assume Laila is worse off than Murray in outcome x ; y is identical to x , except that Laila’s well-being increases by Δw ; y' is also identical to x , except now it is Murray’s well-being that increases by Δw . Then, because of the concavity of $g(\cdot)$, a prioritarian SWF prefers y to y' .

The axiomatic expression of priority for the worse-off is the *Pigou-Dalton* axiom: A pure, gap-diminishing transfer of well-being from someone better off to someone worse off, leaving everyone else unaffected, is an ethical improvement. That is, let $w_l < w_h$, and let $0 < \Delta w < w_h - w_l$. Then if $(w_1, \dots, w_N)$ and $(w_1^*, \dots, w_N^*)$ are such that $w_l^* = w_l + \Delta w$, $w_h^* = w_h - \Delta w$ and everyone else is unaffected, the second vector is better than the first.

If an SWF satisfies Monotonicity, Anonymity, Person Separability, Continuity and Pigou-Dalton, it must take the form $\sum_{i=1}^N g(w_i)$. The utilitarian SWF, $\sum_{i=1}^N w_i$ satisfies Monotonicity, Anonymity, Person Separability, and Continuity, *but not* Pigou-Dalton. In the example immediately above, the utilitarian SWF is indifferent between giving the benefit to Laila and giving it to Murray.

The Pigou-Dalton axiom has much intuitive force, but there is a downside. The utilitarian SWF, when developed into a methodology for choice under uncertainty, readily satisfies the ex ante Pareto principles. The prioritarian SWF does not. So-called “ex ante prioritarianism” satisfies those principles, but at the cost of violating other quite compelling axioms for choice under uncertainty. So-called “ex post prioritarianism” satisfies the latter axioms, but violates the ex ante Pareto principles. In effect, the choice between the utilitarian SWF and prioritarian SWFs comes down to a choice between the ex ante Pareto principles and the Pigou-Dalton axiom. This is a thorny topic that

can't be further discussed here; for more, see Adler (2012: ch. 7, forthcoming) and Mongin and Pivato (2016).

Whereas utilitarianism is a specific SWF, prioritarianism is a *family* of SWFs. There are, after all, *many* different types of $g(\cdot)$ functions. In practice, two are most widely used. The “Atkinson” subfamily of prioritarian SWFs uses the specific $g(\cdot)$ function $g(w_i) = (1 - \gamma)^{-1} w_i^{1-\gamma}$. The “Kolm-Pollak” subfamily uses the function $g(w_i) = -e^{-\lambda w_i}$. Each is parameterized by a single parameter (γ for Atkinson SWFs, λ for Kolm-Pollak SWFs), which captures the degree of priority for the worse-off.

19.6 CONCLUSION

CBA and the SWF framework are powerful tools for assessing governmental policies. Each begins with a methodology for ranking outcomes. In each case, this generalizes to a methodology for ranking probability distributions over outcomes—and therewith policies, understood as such distributions—but given space constraints this chapter has focused on the outcome ranking itself.

CBA's formula is to sum monetary equivalents. This produces a ranking of outcomes that is formally well-behaved (complete and transitive) and respects the Pareto principles. The SWF framework uses a well-being measure, $w(\cdot)$, to convert outcomes into lists (vectors) of interpersonally comparable well-being numbers. The SWF itself is a rule for ranking these vectors. The utilitarian SWF and prioritarian family of SWFs are particularly attractive, and correspond to well-developed positions within philosophical welfarism (utilitarianism and prioritarianism). If $w(\cdot)$ is constructed to respect individual preferences and the SWF is “monotonic,” its outcome ranking will—like CBA's—be formally well-behaved and satisfy the Pareto principles.

While both frameworks satisfy the Pareto principles, they diverge in ranking “Pareto-noncomparable” pairs of outcomes, where those principles do not apply—some preferring the first outcome, others the second.

This chapter has explicated the two frameworks and has also surveyed the arguments for CBA. Both CBA and the SWF approach have been elaborated, in great detail, by large bodies of scholarship within economics; but only CBA is actually used by governments. Is this defensible? I believe not. The case for CBA is not ultimately persuasive—especially with the SWF framework “waiting in the wings.”

REFERENCES

- Note:* Numerous sources below are chapters in Matthew D. Adler and Marc Fleurbaey (eds.), *The Oxford Handbook of Well-Being and Public Policy* (New York: Oxford University Press, 2016), henceforth abbreviated as “OHWBPP.”
- Adler, Matthew D. 2012. *Well-Being and Fair Distribution: Beyond Cost-Benefit Analysis*. New York: Oxford University Press.

- Adler, Matthew D. 2016a. "Benefit-Cost Analysis and Distributional Weights: An Overview." *Review of Environmental Economics and Policy* 10: 264–85.
- Adler, Matthew D. 2016b. "Extended Preferences." In *OHWBPP*, pp. 476–517.
- Adler, Matthew D. 2017. "A Better Calculus for Regulators: From Cost-Benefit Analysis to the Social Welfare Function." Working paper. Available at https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2923829.
- Adler, Matthew D. Forthcoming. *Measuring Social Welfare: An Introduction*. New York: Oxford University Press.
- Adler, Matthew D., and Eric A. Posner. 2006. *New Foundations of Cost-Benefit Analysis*. Cambridge, MA: Harvard University Press.
- Blackorby, Charles, Walter Bossert, and David Donaldson. 2005. *Population Issues in Social Choice Theory, Welfare Economics, and Ethics*. Cambridge: Cambridge University Press.
- Boadway, Robin. 2016. "Cost-Benefit Analysis." In *OHWBPP*, pp. 47–81.
- Boadway, Robin, and Neil Bruce. 1984. *Welfare Economics*. Oxford: Basil Blackwell.
- Boardman, Anthony E., David H. Greenberg, Aidan R. Vining, and David L. Weimer. 2011. *Cost-Benefit Analysis: Concepts and Practice*. 4th ed. Upper Saddle River: Prentice-Hall.
- Bossert, Walter, and John A. Weymark. 2004. "Utility in Social Choice." In Salvador Barberà, Peter J. Hammond, and Christian Seidl (eds.), *Handbook of Utility Theory*, vol. 2 (*Extensions*) (Boston: Kluwer Academic), pp. 1099–177.
- Botzen, W.J. Wouter, and Jeroen C.J.M. van den Bergh. 2014. "Specifications of Social Welfare in Economic Studies of Climate Policy: Overview of Criteria and Related Policy Insights." *Environmental and Resource Economics* 58: 1–33.
- Bykvist, Krister. 2016. "Preference-Based Views of Well-Being." In *OHWBPP*, pp. 321–46.
- Champ, Patricia A., Kevin J. Boyle, and Thomas C. Brown. 2003. *A Primer on Nonmarket Valuation*. Dordrecht: Kluwer Academic Publishers.
- Comim, Flavio, Mozaffar Qizilbash, and Sabina Alkire. 2008. *The Capability Approach: Concepts, Measures, and Applications*. Cambridge: Cambridge University Press.
- Cowell, Frank. 2016. "Inequality and Poverty Measures." In *OHWBPP*, pp. 82–125.
- d'Aspremont, Claude, and Louis Gevers. 2002. "Social Welfare Functionals and Interpersonal Comparability." In Kenneth J. Arrow, Amartya K. Sen and Kotaro Suzumura (eds.), *Handbook of Social Choice and Welfare*, vol. 1. (Amsterdam: Elsevier), pp. 459–541.
- Eggleston, Ben, and Dale E. Miller (eds.). 2014. *The Cambridge Companion to Utilitarianism*. Cambridge: Cambridge University Press.
- Fletcher, Guy (ed.). 2016. *The Routledge Handbook of Philosophy of Well-Being*. Milton Park, UK: Routledge.
- Fleurbaey, Marc. 2016. "Equivalent Income." In *OHWBPP*, pp. 453–75.
- Freeman, A. Myrick. 2003. *The Measurement of Environmental and Resource Values: Theory and Methods*. 2nd ed. Washington, DC: Resources for the Future.
- Fujiwara, Daniel, and Paul Dolan. 2016. "Happiness-Based Policy Analysis." In *OHWBPP*, pp. 286–317.
- Gilboa, Itzhak. 2009. *Theory of Decision under Uncertainty*. Cambridge: Cambridge University Press.
- Graham, Carol. 2016. "Subjective Well-Being in Economics." In *OHWBPP*, pp. 424–50.
- Griffin, James. 1986. *Well-Being: Its Meaning, Measurement, and Moral Importance*. Oxford: Clarendon Press.
- Haybron, Daniel M. 2016. "Mental State Approaches to Well-Being." In *OHWBPP*, pp. 347–78.
- Hicks, John R. 1939. "The Foundations of Welfare Economics." *Economic Journal* 49: 696–712.

- Holtug, Nils. 2017. "Prioritarianism." *Oxford Research Encyclopedia of Politics*, at <http://dx.doi.org/10.1093/acrefore/9780190228637.013.232>.
- Hotelling, Harold. 1938. "The General Welfare in Relation to Problems of Taxation and Railway and Utility Rates." *Econometrica* 6: 242–69.
- Hurka, Thomas. 2016. "Objective Goods." In *OHWBPP*, pp. 379–402.
- Just, Richard E., Darrell L. Hueth, and Andrew Schmitz. 2004. *The Welfare Economics of Public Policy: A Practical Approach to Project and Policy Evaluation*. Cheltenham, UK: Edward Elgar.
- Kaldor, Nicholas. 1939. "Welfare Propositions of Economics and Interpersonal Comparisons of Utility." *Economic Journal* 49: 549–52.
- Kaplow, Louis. 1996. "The Optimal Supply of Public Goods and the Distortionary Cost of Taxation." *National Tax Journal* 49: 513–33.
- Kaplow, Louis. 2004. "On the (Ir)Relevance of Distribution and Labor Supply Distortion to Government Policy." *Journal of Economic Perspectives* 18: 159–75.
- Kaplow, Louis. 2008. *The Theory of Taxation and Public Economics*. Princeton, NJ: Princeton University Press.
- Mongin, Philippe, and Claude d'Aspremont. 1998. "Utility Theory and Ethics." In Salvador Barberà, Peter J. Hammond, and Christian Seidl (eds.), *Handbook of Utility Theory, vol. 1 (Principles)* (Dordrecht: Kluwer Academic), pp. 371–481.
- Mongin, Philippe, and Marcus Pivato. 2016. "Social Evaluation under Risk and Uncertainty." In *OHWBPP*, pp. 711–45.
- Nozick, Robert. 1974. *Anarchy, State and Utopia*. New York: Basic Books.
- Parfit, Derek. 1987. *Reasons and Persons*. Revised and corrected paperback ed. Oxford: Clarendon Press.
- Parfit, Derek. 2000. "Equality or Priority?" In Matthew Clayton and Andrew Williams (eds.), *The Ideal of Equality* (Houndmills: Palgrave), pp. 81–125.
- Renda, Andrea. 2011. *Law and Economics in the RIA World*. Cambridge: Intersentia.
- Robbins, Lionel. 1932. *An Essay on the Nature and Significance of Economic Science*. London: Macmillan.
- Sumner, L.W. 1996. *Welfare, Happiness, and Ethics*. Oxford: Clarendon Press.
- Temkin, Larry S. 1993. *Inequality*. New York: Oxford University Press.
- Tuomala, Matti. 2016. *Optimal Redistributive Taxation*. Oxford: Oxford University Press.
- Weymark, John. 2016. "Social Welfare Functions." In *OHWBPP*, pp. 126–59.
- Wiener, Jonathan B. 2013. "The Diffusion of Regulatory Oversight." In Michael A. Livermore and Richard L. Revesz (eds.), *The Globalization of Cost-Benefit Analysis in Environmental Policy* (Oxford: Oxford University Press), pp. 123–41.

CHAPTER 20

THE NORMATIVE ECONOMICS OF SOCIAL RISK

MARC FLEURBAEY

20.1 INTRODUCTION

IN economic theory, a good way of handling risk and uncertainty is to make everything contingent on the occurrence of particular states of the world. Once this is done, uncertainty is no different than space or time as dimensions along which commodities are differentiated. The Arrow-Debreu extension of general equilibrium theory to uncertainty has famously and successfully inaugurated this strategy (Arrow 1970; Debreu 1959), and the classical welfare theorems linking competitive equilibrium and Pareto optimality have been seamlessly extended to uncertainty in this way—even if this has raised awareness about how heroic the underlying assumptions are (in particular about complete markets and perfect information). The theory of fair allocation, which makes a selection of allocations on the basis of fairness and efficiency criteria (Thomson 2011), or proposes rankings of allocations on the basis of such principles (Fleurbaey and Maniquet 2011), can similarly be directly extended from the context of full certainty to the context of uncertainty—with the only restriction that individual preferences may have a specific form such as expected utility.

But things are not so simple. Harsanyi (1955), in a famous theorem, claimed to prove that the combination of very basic rationality and Pareto-efficiency principles, in the context of risk, implies that the social criterion should be linear in individual von Neumann-Morgenstern (vNM) utilities, which is incompatible with most fairness principles. The basic rationality assumptions in the theorem are simply that both individual and social evaluation relies on expected utility (expected social welfare, for the social level). The Pareto-efficiency principle is that if everyone in the population prefers a lottery to

Table 20.1 Utilitarianism vs. fairness

Lottery I	Heads	Tails	Lottery II	Heads	Tails
Ann	0	2	Ann	2	2
Bob	2	0	Bob	0	0

Table 20.2 Do ex post inequalities matter?

Lottery I	Heads	Tails	Lottery III	Heads	Tails
Ann	0	2	Ann	1	1
Bob	2	0	Bob	1	1

another lottery, the social evaluation must concur. A linear social criterion is of course nothing else, formally, than the standard utilitarian social welfare function, possibly augmented with weights on individual utilities. Many authors have embraced this result and happily adopted utilitarianism as the main approach for welfare economics (such as Broome 1991; Hammond 1987; Mirrlees 1982; Blackorby et al. 1996). But should all non-utilitarian approaches, in particular the fairness approach, be abandoned? Critics of utilitarianism have not surrendered.

Early on, Diamond (1967) noticed that utilitarianism neglects *ex ante* fairness, because it is indifferent about distributing chances among people and among states of the world, considering each individual-state pair as a bucket for receiving utility that is only weighted by the probability of the state, independently of whether the individual has got any chance in other states. To illustrate, (non-weighted) utilitarianism is indifferent between the two lotteries of Table 20.1, where the vNM utilities are given.

Indifference between the two lotteries is hard to defend, because Lottery I is obviously fairer. A utilitarian like Harsanyi would retort that preferring Lottery I to Lottery II is irrational, because in every state of the world they give equivalent distributions of utility.

Broome (1991) noticed that both utilitarianism and the approach advocated by Diamond ignore *ex post* inequalities, because they are not able to discriminate between the lotteries in Table 20.2. In this example, Lottery III achieves equality in every state, unlike Lottery I.

These debates have developed into a general opposition between *ex ante* approaches, that think of social welfare as a function of prospects faced by the individuals, for example, as evaluated by their expected utilities:

$$W(EU_1, \dots, EU_n),$$

and *ex post* approaches that first compute social welfare in each possible final consequence before taking an expected value:

$$EW(U_1, \dots, U_n).$$

Harsanyi's utilitarian theorem can be read as saying that social rationality commands adopting both the *ex post* approach (expected utility at the social level) and the *ex ante* approach (in order to satisfy the Pareto principle with respect to individual expected utilities), and it is not surprising that combining the two formulae requires the function W to have linearity properties. Diamond's critique suggests adopting the *ex ante* approach with an inequality-averse W , and this is also what the adaptation of the theory of fair allocation to contingent commodities entails, whereas Broome's critique would be tackled by an *ex post* approach, also combined with inequality aversion.

This debate between utilitarianism, *ex ante* egalitarianism, and *ex post* egalitarianism is still going on, and one should not underestimate its practical importance. Most policy decisions involve unequal risks for different parts of the population, and therefore elicit very different evaluations from the different approaches. For instance, Finkelstein et al. (2013) find that the marginal utility of income decreases under sickness, and conclude that health insurance should be limited, because it transfers money from (healthy) states with high marginal utility to (unhealthy) states with lower marginal utility. This conclusion is valid for an *ex ante* approach but may run counter to an *ex post* approach with sufficient inequality aversion, if sickness lowers not just marginal utility but also the level of utility, therefore justifying giving a greater priority to the sick and disabled.

More generally, it is common in analyses of the welfare state to compare redistribution with an insurance mechanism. The philosopher Ronald Dworkin (2000) has even based his theory of distributive justice on the idea that redistribution should mimic the allocation that a hypothetical insurance market would generate. On such a hypothetical market, everyone could buy an insurance before knowing one's assets, including personal talents that generate inequalities. This appears appealing at first glance, because it involves a "veil of ignorance" that guarantees impartiality. But the insurance model may be biased for thinking about redistribution, because it is based on individuals making decisions to maximize their expected utility (or some non-expected-utility objective), and is therefore attuned to the *ex ante* approach.

Developing a theory of welfare for the context of risk and uncertainty makes it possible to address various questions that this special context raises and are often asked: 1) Should social decisions be more or less risk averse than the average person? 2) Should we try to avoid large catastrophes more than frequent but limited harms with similar expected impact? 3) Should social decisions be ambiguity averse or stick to the expected utility canon?

In this chapter, I first briefly explain Harsanyi's result and its implications in Section 20.2. In Sections 20.3 and 20.4, I examine the various (*ex ante* and *ex post*) alternatives to utilitarianism that have been proposed. This gives the occasion to examine the first question from the previous paragraph. Section 20.5 is devoted to risk aversion at the

social level and the avoidance of “catastrophes,” thereby addressing the second question from the previous paragraph. In Section 20.6, I discuss difficulties raised both by the requirement of subgroup separability, which becomes extremely constraining under uncertainty, and by dropping it, which makes evaluation a daunting holistic task. Section 20.7 is devoted to ambiguity aversion (the third question from the previous paragraph). Finally, in Section 20.8, I discuss veil-of-ignorance arguments, which are not about real risky situations, but which invoke hypothetical risks in order to find answers to distributive problems (as in Dworkin’s approach). A good theory of how to deal with risk sheds light on how risk can be used to answer distributive questions. Section 20.9 concludes by offering thoughts for future research.

I would like to mention that several surveys and syntheses on welfare economics, risk and uncertainty, can supplement this one for interested readers. An expanded version of this chapter can be found in Fleurbaey (2018), where additional issues (heterogeneous beliefs) and applications are discussed (risk classification, health insurance, and value of life). Weymark’s (1991) chapter is a classical and thorough introduction to, and a critical assessment of, Harsanyi’s arguments for utilitarianism (the aggregation theorem and the impartial observer argument). Mongin and d’Aspremont (1998) develop a detailed and deep discussion of utilitarianism, including the uncertainty context, with a thorough analysis of welfarism in the background. Adler and Sanichirico (2006) offer a very detailed discussion of the *ex ante* versus *ex post* debate and show, through interesting examples, that the *ex post* approach plays an important role in legal theory. Hammond (1982) gives a good summary of justifications of utilitarianism under uncertainty and a thoughtful discussion of the role of information in welfare evaluation. Hild et al. (2008) contains a thought-provoking analysis of Harsanyi’s theorem and the problem of aggregating beliefs and preferences, with a comparison of the *ex ante* and *ex post* approaches. Fleurbaey (2008: ch. 6) studies responsibility for risk-taking and luck, with applications to insurance markets and Dworkin’s hypothetical insurance theory. Finally, Mongin and Pivato (2016) discuss many of the topics of this chapter, with a focus on the *ex ante* vs *ex post* debate, at a more abstract level.

20.2 HARSANYI’S THEOREM

As recalled in the introduction, Harsanyi’s (1955) theorem derives strong conclusions about the linearity of the social preferences from apparently compelling rationality and efficiency requirements. Harsanyi’s theorem can be formulated in this way:

Theorem: If individual preferences and social preferences are of the expected utility (EU) type (with objective probabilities) and satisfy the Pareto principle, then social preferences can be represented by a vNM utility that is a weighted sum of individual vNM utilities.

For Harsanyi, this was a decisive argument in favor of utilitarianism. For Sen (1974, 1977), however, it was no more than a representation result, with little substantial connection to utilitarianism. Weymark (1991) developed a more precise criticism, arguing that because utilitarianism need not rely on individual vNM utilities, the (weighted) sum of vNM utilities is not necessarily linear in the utilities that utilitarianism would use.

This critique, nevertheless, ignores the fact that the axioms of Harsanyi's theorem appear to be endorsed by utilitarians. Social rationality and the Pareto principle are popular among them. Recognizing this fact, Fleurbaey and Mongin (2016) offer the following variant of Harsanyi's theorem that shows the formal link established with utilitarianism through the theorem:

Theorem: If utilitarian social preferences on sure options, based on a specific sum $\sum_i u_i^*$ of particular "utilitarian" utilities u_i^* , can be extended into social preferences over state-contingent prospects that satisfy the Pareto principle and the EU property, then any social preference over state-contingent prospects satisfying Pareto and the EU property can be represented on sure options by a weighted sum of the utilitarian utilities u_i^* .

Even if the connection between Harsanyi's theorem and utilitarianism appears substantial, this leaves open the question of whether the theorem is an argument for the linearity of social preferences or an impossibility theorem questioning the validity of its axioms. Here are critical concerns about the implications of linear social preferences.

First, in the context of a simple model with only one commodity (money), and in which individuals have identical risk attitudes, social preferences that are linear in vNM utilities exhibit the property that the common risk aversion in individual utilities determines inequality aversion over the commodity at the social level. This is especially clear if the same vNM utility function $u(x)$ is chosen to represent all individuals' preferences, and equal weights are used in the social welfare function, so that $\sum_{i=1}^n u(x_i)$ is the criterion used to evaluate the allocation $(x_1, \dots, x_n)$. This link between inequality aversion and risk aversion is strange because the former has to do with the ethical issue of distributing between people, whereas the latter has to do with the prudential issue of distributing among states of the world for any given person. Why should the social evaluator be inequality neutral when the population is risk neutral, for instance?

Second, in the same simple model but allowing for heterogeneous risk attitudes, it appears strange that the social marginal utility of money remains constant for risk neutral individuals whereas it declines quickly for strongly risk averse individuals. This means in particular that fairness principles in terms of reduction of resource inequality (such as the Pigou-Dalton transfer principle) cannot be accommodated in such a setting.

Therefore, one can read Harsanyi's theorem as revealing a tension, that is specific to the context of risk and uncertainty, between rationality (embodied in the EU property at the individual and the social level), Pareto (in particular, respecting individual risk attitudes when taking risks), and equity (such as the Pigou-Dalton principle). The

utilitarian option is weak on the equity side, and in the next section we explore alternatives which are stronger on equity but relax one of the other two pillars.

20.3 *EX ANTE* EGALITARIANISM

Dropping the EU property at the social level has been quite popular since Diamond's (1967) critique attracted interest for *ex ante* egalitarianism. Machina (1989) also endorsed the idea that the EU property is not compelling for decisions affecting different people. Epstein and Segal (1992) developed an axiomatic justification of *ex ante* egalitarianism based on a social preference for randomizing.

Ex ante egalitarianism is popular in economics because, as explained in the introduction, it is the most natural extension of inequality averse social welfare functions to the context of risk. However, it has already been explained that *ex ante* egalitarianism triggers violations of statewise dominance, that is, the principle that if a lottery implies a better outcome than another in every possible state of the world, it must be preferred. Lotteries I and II in the introduction yield equally good consequences in every state, therefore are equally good by statewise dominance. The literature has actually shown that one can replace the EU condition at the social level by the weaker statewise dominance condition and still obtain Harsanyi's theorem (Blackorby et al. 2004; Gajdos et al. 2008; Fleurbaey 2009; Chambers and Hayashi 2014). Mongin and Pivato (2015) and Zuber (2016) provide the most general analysis of this type of result.

One may be of two minds about this issue. One view is that statewise dominance being the most basic rationality condition under uncertainty, it is very troubling that *ex ante* egalitarianism systematically violates it. The opposite view is that preferring Lottery I to Lottery II is very sensible, and that applying statewise dominance appears narrow-minded in this type of case. When comparing the two lotteries in state "Heads," one should not forget what happens in state "Tails," and statewise dominance appears to shed relevant counterfactual information.

The obvious way to reconcile the pull of statewise dominance with the relevance of counterfactual information is to introduce the relevant counterfactual information in the description of final consequences, as in the revised description of Lotteries I and II in Table 20.3.

The notation u/v means that vNM utility is u in the current state and would have been v in the other state. Given that one can rationally prefer the distribution $(0/2, 2/0)$ to the

Table 20.3 Enriched description of consequences

Lottery I	Heads	Tails	Lottery II	Heads	Tails
Ann	0/2	2/0	Ann	2/2	2/2
Bob	2/0	0/2	Bob	0/0	0/0

distribution $(2/2, 0/0)$, because the former involves “fair chances,” it is now compatible with statewise dominance to prefer Lottery I. Note that a very slight preference for $(0/2, 2/0)$ over $(2/2, 0/0)$ suffices to get the desired preference for Lottery I. It is not necessary to evaluate lotteries solely in terms of individual expected utilities and it is not necessary to consider that Lottery I achieves perfect equality.

Let us take stock. A crude application of the *ex ante* approach appears to run into serious rationality troubles. Such problems can be avoided by keeping an approach that satisfies dominance and refining the description of final consequences in order to take account of relevant counterfactual information and claims linked to past history. The important lesson here is that, while the *ex ante* approach focuses exclusively on *ex ante* individual prospects (as evaluated by expected utilities), it appears sufficient to have the slightest concern for past chances in order to obtain the preferences for randomizing that Diamond as well as Epstein and Segal advocate.

20.4 EX POST EGALITARIANISM

Solving the irrationality problems of *ex ante* egalitarianism by enriching the description of final consequences and respecting statewise dominance, or even the expected utility property at the social level, means adopting the *ex post* approach. However, unlike the *ex ante* approach, the *ex post* approach has difficulties combining inequality aversion with the Pareto principle. The comparison of Lotteries I and III, in the introduction, provides an example in which all individuals are willing to take a risk (because Lottery I appears just as good to each of them), but at the social level this appears undesirable because the gamble produces inequalities in the final consequences.

The reason why the Pareto principle clashes with a concern for inequalities in the final allocations is straightforward. Just as statewise dominance makes it impossible to keep track of counterfactuals (in the non-enriched framework), the Pareto principle makes it impossible to keep track of correlations between individuals and of inequalities in particular states.

Moreover, the clash between Pareto and inequality aversion, under statewise dominance, is deep and cannot be fixed by enriching information. Let us rewrite Lotteries I and III by enriching the description of final consequences, specifying the counterfactuals (as in Table 20.4).

Table 20.4 Pareto vs. inequality aversion

Lottery I	Heads	Tails	Lottery III	Heads	Tails
Ann	0/2	2/0	Ann	1/1	1/1
Bob	2/0	0/2	Bob	1/1	1/1

In state “Heads,” one has to compare $(0/2, 2/0)$ to $(1/1, 1/1)$, and a similar comparison must be made in state “Tails.” Now, Pareto would require indifference between the two lotteries, and under statewise dominance (and given the symmetry between final consequences in the two states), the only way to obtain such indifference is to be indifferent between the final consequences $(0/2, 2/0)$ to $(1/1, 1/1)$. The rich information provided here makes it possible to see that both individuals have equal expected utilities in the two distributions and are therefore indifferent. But clearly, the former distribution is more unequal than the latter. To declare them both perfectly equal would completely ignore the unequal distribution of luck in one of them. Even if a wager is fair, it seems very strange to say that the players are equally well-off *ex post*. Therefore, no matter how weak the concern for inequality is, it runs against the Pareto principle.

It was possible to reconcile statewise dominance with a concern for *ex ante* fairness by enriching the description of final consequences. But it appears impossible to reconcile the Pareto principle with a concern for inequality (unless one takes the farfetched view that unequal luck does not generate inequality). This, of course, does not imply that the Pareto principle should be dropped entirely in the presence of a concern for inequality. The literature has examined several weakenings of the Pareto principle for the context of risk.

Bommier and Zuber (2008) propose to restrict Pareto to statistically independent prospects, and characterize a class of social welfare functions that includes utilitarianism and additional multiplicative functions. Fleurbaey (2010) proposes an axiom of “Pareto for equal risk” that restricts Pareto to lotteries that preserve equality of utilities in every state, and obtains a social welfare function such as:

$$EW(U_1, \dots, U_n),$$

where W is normalized so that $W(x, \dots, x) = x$. Such a function is called an equally distributed equivalent (EDE) because by construction, $W(U_1, \dots, U_n)$ is the solution x to the equation

$$W(U_1, \dots, U_n) = W(x, \dots, x).$$

This criterion boils down to EU_i in case of perfect correlation, hence respecting the Pareto principle in this case. Fleurbaey also considers a stronger Pareto condition that applies to risks in which there is no reranking of individuals’ utilities across states, and obtains a specification of the W function that is rank-dependent like the Gini social welfare function. In another variant, the Pareto principle is applied when a subgroup does not bear any risk and is indifferent between two lotteries, whereas the rest of the population is submitted to risk but does not suffer any inequality. However, this last condition is strong enough to preserve the Harsanyi result.

In an economic model allowing for interpersonal comparisons made with other utilities than the vNM, Fleurbaey and Zuber (2017) rely on another variant, “Pareto for equal or no risk,” which applies Pareto when the two lotteries to be compared

belong to the union of the set of lotteries which preserve equality in every state and the set of sure allocations. They obtain a social criterion that assesses the (riskless) distribution of certainty-equivalent bundles of the EDE lottery.

Gajdos and Maurin (2004) restrict Pareto to *ex post* Pareto (that is, when there is a Pareto dominance in every state) and advocate social welfare functions like:

$$\alpha EW(U_1, \dots, U_n) + (1-\alpha) W(EU_1, \dots, EU_n),$$

which combine the *ex ante* and the *ex post* approach in a convex combination. Note, however, that this type of social welfare function can respect a stronger Pareto condition if the first term itself satisfies such a stronger condition.

This diversity of Pareto conditions is bewildering. How to select the proper scope of the Pareto principle in this context? A useful guide is to notice that statewise dominance does not leave any room for Pareto when the final consequences are all equally good for the social criterion. Even if individuals bear risk, there is no risk at the social level, making it questionable to seek to respect individual risk attitudes.

Let us illustrate this idea with a more concrete example. Consider the comparison between Lotteries I and III again, and assume that the world is deterministic. This means that the identity of the winner in Lottery I has already been determined. For instance, the name of the winner is written in a sealed envelope. An observer here can hold two reasonings. The *ex ante* reasoning goes like this: neither Ann nor Bob knows if (s)he is the winner, they have the same expected utility. The *ex post* reasoning goes like this: one of them is the winner, with true utility 2; the other loses, with true utility 0; the identity of the winner is unknown but for social evaluation this is irrelevant information. In this example it is clear that the *ex ante* reasoning is nurturing the illusion that both have chances, whereas the truth of the matter is better captured by the *ex post* reasoning. There is in fact no risk, only imperfect information, and fortunately, the missing information is irrelevant for the evaluation. Individual risk attitudes in this context are not relevant because individuals do not truly bear risk.

This example may suggest that the application of the Pareto principle could depend on metaphysical questions about determinism. But it seems reasonable to argue that, independently of whether determinism is true or not, the case in which there is no risk at the social level is the case in which the Pareto principle is the least compelling. In such a case it is possible to take account of the distribution of individuals' final preferences on actual payoffs, and it seems dubious to insist on respecting their *ex ante* preferences that ignore their final payoffs when the distribution of their final preferences is already known. In the classical understanding of consumer sovereignty, fully informed preferences take precedence over less informed preferences.

When does it happen that there is no risk at the social level? When the final distribution is already known *ex ante* and only the identity of the individuals positioned at different ranks in the distribution is unknown. This happens when there is negative correlation between individuals' payoffs, but one also comes close to it when payoffs are statistically independent and the population is large. Therefore Bommier and

Zuber's (2008) Pareto condition, which applies to statistically independent risks, is not the most compelling. The case that is the farthest from the absence of risk at the social level is not the case of independent risks but the case in which equality is preserved in every state, so that social risk is aligned with every individual's risk. This is the case in which social risk is maximal for a given degree of individual risk.

We can therefore consider the axiom of Pareto for equal risk as a candidate for the most compelling Pareto axiom under risk. Stronger Pareto conditions will help specify the social welfare function further. Note that Fleurbaey and Zuber (2017) show that their axiom of Pareto for equal or no risk may already be going too far, as it is incompatible with the expected utility property unless the social criterion gives absolute priority to some individuals (such as the worst-off, as in the maximin approach).

Let us take stock again. Harsanyi's theorem shows that combining social rationality and Pareto puts severe constraints on inequality aversion. While the *ex ante* approach seems problematic because it violates rationality and ignores inequalities due to unequal luck, the *ex post* approach is able to preserve rationality and sustain inequality aversion at the reasonable cost of ignoring individual risk attitudes when they seem the least relevant, that is, when there is no risk at the social level so that it is actually possible to respect fully informed individual preferences rather than *ex ante* preferences.

20.5 CATASTROPHE AVOIDANCE VERSUS EQUITY

In a seminal paper, Keeney (1980) studied the specific problem of allocating fatality risk between individuals, assuming independent risks, and noticed an interesting clash. Risk aversion over the number of casualties, an attitude he called "catastrophe avoidance," implies a preference for concentrating risks over "suicide patrols," in contradiction with *ex ante* equity which prefers spreading risk over more people.

For instance, consider a given expected number of casualties, say, one for a given population of n individuals. The variance of this risk is zero if exactly one person is sacrificed for sure. The variance is maximal if everyone in the population has a $1/n$ chance of accident. Risk aversion at the social level therefore implies inequity, and conversely, a preference for an equal distribution of risks (*ex ante* inequality aversion) requires a risk-loving attitude at the social level.

Keeney restricted attention to independent risks, but catastrophe avoidance also implies a preference for inverse correlation over positive correlation, because an inverse correlation stabilizes the distribution of final positions, whereas a positive correlation makes all boats float on the same tide. In this case, catastrophe avoidance goes against *ex post* inequality aversion. Consider again the example with an expected casualty of one in n individuals. With inverse correlation it is possible to make sure

exactly one individual is hit, while preserving *ex ante* equity because *ex ante* everyone can have a $1/n$ chance of accident. With perfect correlation, everyone still has a $1/n$ chance of accident, but the probability of the whole population being killed is maximal, at $1/n$. Yet, this is what *ex post* inequality aversion recommends to prefer because equality if guaranteed *ex post* (and the probability that all will be safe is also maximal).

In conclusion, catastrophe avoidance runs against both *ex ante* and *ex post* equity. It is therefore puzzling that it often appears a reasonable and attractive attitude. One possible explanation for its intuitive appeal is that catastrophes have a multiplier effect through externalities in society. If a family dies, this is very sad but does not disrupt society much. If the equivalent of a whole city dies in a country, life is completely disrupted for most people. Therefore, in real life externalities do make catastrophes much bigger than in theoretical models. As an extreme example, observe that wiping out a whole generation would prevent future generations from existing.

Another possible explanation of the appeal of catastrophe avoidance is that risk aversion for society seems a natural extension of risk aversion at the individual level. This may be, in part, a mathematical illusion. Indeed, inequality aversion imposes extra concavity on individual utilities, but if one wants to respect the Pareto principle in case of perfect correlation, this extra concavity needs to be compensated by a convex transform, implying a risk-loving attitude at the social level. Consider for instance the case of a standard social welfare function $\sum_i \varphi(u_i)$, where u_i is i 's vNM function and φ is a concave function embodying inequality aversion. The EDE in this case is

$$\varphi^{-1}\left(\frac{1}{n}\sum_i \varphi(u_i)\right),$$

so that taking the expected value of the EDE as the social criterion may be read as applying the convex, risk-loving transform φ^{-1} to the value of social welfare. However, Ferranna (2017) shows that in an economic model, the social welfare function may take the expected value of an EDE defined in terms of resources rather than utility, and then catastrophe avoidance arises when risk aversion in the expected value is greater than inequality aversion in the EDE.

A very detailed review of the literature on catastrophic risk and of attitudes about catastrophes by Rheinberger and Treich (2015) recently concludes that catastrophe avoidance is not that popular after all and, among many possible behavioral explanations, highlights equity as the main factor that seems to be pulling people and decision-makers away from it.

There is a related literature on the precautionary principle (for example, Gollier et al. 2000) that relies on a model with a single risk-averse decision-maker and identifies the link between risk aversion, irreversibility of damages, and learning. It would be interesting to extend this literature to social criteria that are sensitive to the distribution of well-being, because the precautionary principle is likely to be modified if the potential damages are concentrated on a particular population.

20.6 THE SEPARABILITY CONUNDRUM

Separability (also called subgroup separability) is a property of social evaluation that makes it possible to ignore unaffected populations when evaluating projects that do not impact the whole population. It is satisfied by utilitarianism thanks to its additive form, and also by the *ex ante* approach when it involves an additively separable social welfare function such as $\sum_i \varphi(EU_i)$. Separability is practically very convenient because it makes it possible to focus on the affected population. It is a core principle of prioritarianism as defended by Parfit (1991) and Adler (2013).

Non-separable social criteria are problematic because in principle they require taking account of the whole distribution of well-being over the whole relevant population (potentially the whole human population from its poorly known origins to its uncertain end). It is of course possible to apply a non-separable social welfare criterion to a subpopulation, but if the subpopulation changes between evaluations, inconsistent results are likely to occur. For instance, ignoring past generations with a non-separable criterion typically generates time inconsistency when interim generations are dropped in a future re-evaluation of a plan.

It turns out that *ex post* egalitarianism implies non-separability. Consider the comparison between Lotteries I and IV in Table 20.5.

Bob is not affected at all by the change from one lottery to the other. Under separability, one should ignore the particulars of his fixed situation, and, for instance, make the same evaluation for the lotteries shown in Table 20.6.

This means that if Lottery IV is preferred to Lottery I (because it offers equality *ex post* in every state), one should also, by separability, prefer Lottery I' to Lottery IV', which appears absurd.

It is clear from this example that separability requires ignoring the correlation between people's payoffs, and this is similar to what the full Pareto condition implies. Even the weak separability condition that applies only to individuals bearing no risk (in addition to being unaffected) is strong enough to bring Harsanyi's theorem back, under Pareto for equal risk and statewise dominance (Fleurbaey 2010: Th. 4). To get an intuition for this result, consider the lotteries in Table 20.7.

By Pareto for equal risk applied to Ann (ignoring the presence of Bob by separability), Lotteries III and V are equally good. Similarly, by Pareto for equal risk applied to Ann and Bob, Lotteries III and VI are equally good. By transitivity, Lotteries V and VI are

Table 20.5 Ex post egalitarianism vs. separability

Lottery I	Heads	Tails	Lottery IV	Heads	Tails
Ann	0	2	Ann	2	0
Bob	2	0	Bob	2	0

Table 20.6 Implications of separability

Lottery IV'	Heads	Tails	Lottery I'	Heads	Tails
Ann	0	2	Ann	2	0
Bob	0	2	Bob	0	2

Table 20.7 Pareto for equal risk, separability and statewise dominance vs. inequality aversion

Lottery III	Heads	Tails	Lottery V	Heads	Tails	Lottery VI	Heads	Tails
Ann	1	1	Ann	2	0	Ann	1.5	0.5
Bob	1	1	Bob	1	1	Bob	1.5	0.5

equally good. By statewise dominance, Lottery VI would be strictly preferred to Lottery V if equalizing utilities were considered a strict improvement. Therefore, indifference between the two lotteries is incompatible with inequality aversion.

The dilemma revealed by this refinement of Harsanyi's theorem is therefore involving four instead of three conditions: Pareto, social rationality, equity, and separability. Utilitarianism relaxes equity, *ex ante* egalitarianism relaxes rationality, and *ex post* egalitarianism weakens both Pareto and separability.

Let us revisit the *ex post* approach. Is dropping separability really a nightmare for practical evaluations? There is actually some attraction to the idea that comprehensive evaluation should take account of the whole context of human history. The greatest statesmen (and women) generally put their action in the perspective of the long-term human history, including the past, and this may be more than pure rhetorics. As explained in Fleurbaey (2018) and Ferranna (2017), a non-separable evaluation proceeds like a separable evaluation except for the presence of "background parameters" that operate like distortions of probabilities of states of the world, making the evaluation optimistic or pessimistic depending on the past history.

20.7 AMBIGUITY AVERSION

So far we have assumed that expected utility was the criterion of choice both at the individual and social level. But other non-EU criteria have been studied in decision theory. Here we will focus on the context of uncertainty in which probabilities are not well defined. There are two salient options here: Either the evaluator should construct probabilities first and then apply the EU criterion, following a method recommended by Savage (1972), or

the evaluator should adopt an ambiguity-averse criterion, such as the maxmin EU $\min_{p \in P} E_p U$ (Gilboa and Schmeidler 1989), where P is the set of probabilities considered and E_p denotes the expected value operator relying on probability vector p .

An interesting debate bears on the normative value of ambiguity aversion. Al-Najjar and Weinstein (2009a, 2009b) have severely criticized ambiguity-averse approaches, pinpointing the irrationalities they typically induce (violations of dominance and time consistency, as well as refusal of free information). On the contrary, Gilboa et al. (2009) argue for a more flexible notion of rationality, and Heal and Millner (2015), focusing on the example of climate policy (for which there are many scientific uncertainties), argue that it is dishonest to construct precise probabilities when there is no consensus on probabilities, and recommend ambiguity aversion in social evaluation. For individual preferences, in contrast, Heal and Millner consider the EU criterion acceptable, because it is more acceptable for a single individual to make the leap of adopting specific probabilities.

Taking inspiration from Al-Najjar and Weinstein (2009a, 2009b) and Fleurbaey (2016), I would like to defend the normative standing of the EU approach because of the disturbing irrational implications of ambiguity aversion. I will also argue that ambiguity aversion is *less* acceptable for social evaluation than for individual evaluation, because in the latter context it is actually compatible with the EU approach.

Here is an example of policy choice that illustrates the strange implications of ambiguity aversion. Suppose that a pandemic is starting and that there are two possible types, A or B. There are four policy options. One is a costly general-purpose *vaccination* campaign. Another is an equally costly but slightly more effective targeted vaccination campaign that requires knowing the type of pandemic. The third and fourth options are less costly general-purpose and targeted *prevention* campaigns, the latter also requiring knowledge of the pandemic type. There is no reason to believe that prevention is less efficient than vaccination, whatever the type of pandemic, but its probability of success is less well known because it depends on the population's behavioral reactions. Moreover, the prevention policy works better for one of the pandemic types but it is not known which, even when it is targeted. This is because the mechanisms by which prevention works are less well understood than for the vaccines. To fix ideas, assume that the probabilities of success are as in Table 20.8.

Assume that the two types of pandemic are considered equally likely. If the type is not known, the probability of success under general prevention is estimated at $0.5 \times p + 0.5 \times (176\% - p) = 88\%$ percent, and because it is less costly than either vaccination, it is the

Table 20.8 Prevention vs. vaccination

Pandemic type	General vaccine	Targeted vaccine	General prevention	Targeted prevention
A	86%	88%	$78\% < p < 98\%$	$80\% < p < 100\%$
B	86%	88%	$176\% - p$	$180\% - p$

preferred option. But if the type is known, the probability of success under (targeted) prevention is considered ambiguous, possibly as low as 80 percent, and then it may appear preferable to resort to targeted vaccination due to ambiguity aversion. It may even happen that the general vaccine is also preferred to targeted prevention!

This sort of social preferences is obviously irrational. There is no reason to believe that prevention works less well, and it is less costly. If offered to test the type of pandemic, the decision-maker would refuse the test, fearing that it would induce her to switch to the costly (targeted) vaccination with no greater chance of success. Such refusal of free and useful information is hard to justify in a context in which information induces no strategic disadvantage. Another related irrational implication is time inconsistency. Indeed, planning ahead of the test, the best plan is to perform targeted prevention whatever the type of pandemic, but once the test is done and the type is known, the temptation is to switch to vaccination.

This example may look farfetched, but it should make clear that ambiguity aversion systematically induces a preference for more costly, less effective, but less ambiguous options (which may perhaps contribute to a bias of policymakers for technical fixes over behavioral policies, if there is more ambiguity in behavioral responses), and a preference for ignorance or useless randomization when this reduces ambiguity.

Ambiguity aversion is irrational and costly for social preferences, but I would like to argue that it is admissible for individual preferences, when it is associated with negative feelings which have an impact on well-being. It is actually possible for an individual to satisfy the EU property and behave with apparent ambiguity aversion, just because of the emotions triggered by ambiguity. Consider an individual facing Ellsberg's urns (Ellsberg 1961). Urn I has a 50-50 proportion of white and blue balls, and Urn II has red and yellow balls in unknown proportions. Betting on white or blue does not trigger as much anxiety as betting on red and yellow, and therefore, taking account of emotions in addition to monetary rewards, an EU maximizer can rationally prefer to avoid ambiguity. In contrast, for social preferences it would be silly to let the evaluation be influenced by the evaluator's anxiety about ambiguity. Of course, if ambiguity triggers emotions in the population, the social evaluator may rationally want to take account of them, but this is very different from adopting an ambiguity-averse decision criterion. It should be mentioned that there is empirical work questioning the prevalence of ambiguity aversion in the population, especially for losses as opposed to gains (see conflicting results in Kocher et al. 2015 and Voorhoeve et al. 2016). Bradley (2016) theorizes a different type of attitude, in which people attribute value to chances (over and beyond final outcomes) in such a way as to generate ambiguity aversion for gains and ambiguity love for losses.

In conclusion, the EU criterion seems unshaken by ambiguity as the most rational approach to decision-making under uncertainty, even if it requires making the leap of constructing precise probabilities. For individual decision-makers, it is nevertheless possible for emotions to induce behavior that resembles ambiguity-averse criteria such as maxmin EU. For social preferences, it would be hard to make sense of the evaluator's emotions, such as ambiguity anxiety, playing a role in the evaluation of the fate of society.

20.8 VEIL-OF-IGNORANCE ARGUMENTS

While this chapter is mainly about how to deal with actual risk and uncertainty in social evaluation, it is natural here to also discuss the interesting fact that welfare economics has often relied on risk as a thought experiment in order to derive conclusions about non-risky distributional issues.

Harsanyi (1953, 1977) proposed to imagine what an impartial observer would prefer if she could end up in everyone's shoes, as a guide to construct social preferences. An observer using the EU criterion would maximize her expected utility, which would, in this special lottery, assuming that there is equal chance of becoming anyone, coincide with the average utility of the population. Of course, for this to make sense or even to "justify" utilitarianism, one must assume that the observer's vNM utility would coincide, if she happened to become Mr./Ms. i , with i 's utility as one would like to measure it.

A large literature has explored this impartial-observer argument and questioned the various assumptions: the application of equal probabilities (Karni 1998; Gajdos and Kandil 2008), the application of the EU criterion and the definition of preferences over identity lotteries (Karni and Weymark 1998; Grant et al. 2010, 2011, 2012), and the assumption that the observer's utilities can be linked to the population's utilities (Pattanaik 1968; Weymark 1991). Useful syntheses on this literature can be found in Mongin (2001) and Mongin and Pivato (2016).

Similar arguments have been popular in philosophy. Rawls (1971) has described an "original position" in which, under a "veil of ignorance" hiding their true identities and characteristics, the members of society should design the basic institutions of society before falling back into their own shoes and living within these institutions, normally pursuing their personal interests. Dworkin (2000), as mentioned earlier, designed his theory of egalitarianism around the model of a hypothetical insurance market in which, before being born but already endowed with their adult risk attitudes, and endowed with equal wealth, individuals would take insurance contracts that promise to compensate them for bad luck in terms of personal talents, social background, and similar life circumstances.

Rather than examining this literature in detail, I want to focus on a key objection to the impartial observer argument raised by Kolm (1996) and Roemer (1996). The veil of ignorance about identities and characteristics such as talent does guarantee impartiality of preferences. But impartiality is only the first pillar of a theory of justice and does not at all guarantee fairness. Consider for instance the class of social welfare functions $\sum_i \varphi(u_i)$. They are all impartial (with respect of the distribution of utilities as measured by u_i) but cannot all be fair. It is possible to be impartial and to favor the best-off, and this indeed can happen with the impartial-observer argument if the observer is risk-loving. Even if an assumption of equality between the observer's utility in i 's shoes and i 's utility may prevent this from happening, it does not at all guarantee that the observer favors the worst-off as much as fairness requires. On the contrary, it is likely that under a veil of ignorance, an observer is willing to take risks for herself that

should not inspire distributional decisions between people. Personal trade-offs between different possible situations for oneself seem an inadequate inspiration to take resources away from a disadvantaged person and transfer them to better-off people.

It is interesting to note that an empirical literature eliciting individual opinions about veil-of-ignorance schemes through questionnaires has often found that respondents exhibit a stronger inequality aversion when making a decision about redistribution for other people than the risk aversion they have when considering risky situations in which they are involved (see, in particular, Herne and Suojanen 2004; Traub et al. 2005; Amiel et al. 2009). This pattern is not very strong and there are differences across countries, but such results suggest that for ordinary people, the leap from risk aversion to inequality aversion that veil-of-ignorance arguments advocate is far from natural.

Summing up, the veil of ignorance, which is quite popular in economics as well as in philosophy, does not deserve its popularity. It guarantees impartiality, but does not guarantee fairness and is actually likely to be biased against the disadvantaged because personal trade-offs between states of the world, if translated into distributional transfers, would take away from the disadvantaged too easily.

20.9 CONCLUSION

The welfare economics of risk and uncertainty is still very much a moving field that has provided an interesting variety of results and arguments. Let us try to summarize the path covered in this chapter in a few key points:

1. Harsanyi's theorem, progressively refined in several variants, uncovers a deep tension between the Pareto principle (respecting risk attitudes); social rationality (the EU property or even simply statewise dominance); equity (giving substantial priority to the worst-off); and separability (ignoring unaffected populations when evaluating a project). Utilitarianism, in this dilemma, consists of relaxing equity.

2. The most natural extension of non-utilitarian standard welfare criteria (either social welfare functions, or fair allocation rules and social orderings) to risk, which relies on state-contingent commodities, involves dropping rationality in the above four-fold dilemma and therefore may not be advisable.

3. The Pareto principle is a weak link in the dilemma, because one should beware of spurious unanimities (of preferences or even of beliefs) that are not robust to learning and to information-sharing. The most compelling form of Pareto is Pareto for equal risk, which implies, under statewise dominance, that one should focus on EDE prospects (which substitute an EDE allocation for every final consequence in every state). Inequality aversion in the EDE formula means that social evaluation takes account of unequal luck in addition to other sources of inequalities.

4. Separability is another weak link in the dilemma, because it forces to ignore correlations between individual payoffs and inequalities in final consequences. A non-separable evaluation seems a daunting task if it involves looking at the big picture of

human history for every project evaluation, but it only boils down to introducing a few background parameters in relative priorities across states of the world.

5. Catastrophe avoidance and risk aversion at the social level actually run against *ex ante* and *ex post* equity, requiring caution when thinking of applying such attitudes to policy contexts.

6. Ambiguity aversion is a recipe for wasteful and irrational decisions that seem unjustifiable in the context of social decision-making (unlike individual decision-making, where anxiety about ambiguity may trigger feelings that even an EU maximizer may want to take into account).

7. Veil-of-ignorance arguments are unduly popular because they guarantee impartiality but do not guarantee fairness, and are actually prone to unfairness against the disadvantaged individuals with low marginal utility.

There are many questions that this chapter has not discussed and that, in general, deserve attention in future research. Here is a sample:

1. Interesting social criteria, in the context of riskless allocations, rely on interpersonal comparisons that are not based on vNM utilities. Extending such criteria to the context of risk involves more complex analysis than what has been explained in this chapter. As analyzed in Fleurbaey and Zuber (2017), one then obtains a social criterion that is a weighted sum of individual vNM utilities applied to EDE prospects, where the EDE formula relies on the non-vNM interpersonal comparisons. Selecting the weights in this sum is an interesting problem.

2. Determining the class of appropriate Pareto axioms for the context of risk and uncertainty is still very much work in progress. Even if one agrees with the (of course debatable) view that Pareto for equal risk is the most compelling axiom, there are many extensions of it that one can think of, and the compatibility of such extensions with the other principles in the four-fold dilemma highlighted above deserves to be studied thoroughly.

3. Most of the applied theory of insurance and welfare policies relies either on utilitarianism or on *ex ante* egalitarianism, and it is interesting to revisit this corpus with the tools of *ex post* egalitarianism. An emerging literature in this vein shows that: cost-benefit analysis can use approximate mean-variance formulae that take account of *ex post* inequalities (Ferranna 2015a); *ex ante* Pareto improvements in insurance policy may actually harm well-defined categories of individuals *ex post* (Fleurbaey 2010); regulation against risk classification such as genetic testing may reduce or increase *ex post* inequality unambiguously depending on whether a pooling equilibrium or a separating equilibrium with adverse selection ensues and how much coverage is offered in the former (Hoy 2006; Durnin et al. 2012); optimal health insurance involves public coverage with the option to buy additional insurance for a wide class of social welfare functions (Leach 2010); on migration policy, opening borders harms some disadvantaged migrants and may be rejected by *ex post* (global) egalitarianism (Roemer 2006); the social cost of carbon depends on risk aversion, intragenerational inequality aversion, and intergenerational inequality aversion jointly with empirical assumptions about the distribution of growth

and of climate impacts (Ferranna 2015b); inequalities in longevity may deserve more attention than inequalities in life expectancy (Fleurbaey et al. 2014); results about the business-cycle profile of unemployment benefits or about precautions against catastrophic risk may be altered by inequality aversion (Fleurbaey and Zuber 2017); and formulae for discount rates contain additional fairness and correlation terms under *ex post* egalitarianism (Fleurbaey and Zuber 2015).

4. The incorporation of considerations of *ex ante* fairness in the description of final consequences, in order to render *ex ante* fairness compatible with rationality, is proposed in the chapter but no precise way of doing it is articulated and this remains quite a challenge. Tentative explorations of this topic can be found in Fleurbaey et al. (2015) and Voorhoeve and Fleurbaey (2016), but they are no more than first forays into a difficult field.

5. Risk and time are linked, and the longer the horizon, the more difficult the analysis of risk. In the long run, different states of the world involve different populations of different sizes, and therefore it is important to develop a full-fledged welfare economics of population in conjunction with the welfare economics of risk. Population ethics is a difficult field, full of dilemmas (see Asheim and Zuber 2014 for a good summary and a constructive proposal), and the merging of the two fields is an important domain for research (see attempts in Fleurbaey and Zuber 2015; Asheim and Zuber 2016).

6. Finally, although ambiguity aversion appears questionable for social evaluation, there are many cases in which decision-makers have a hard time coming up with precise probabilities, and it would be useful to develop practical tools that would enable decision-makers to approximately find the best options without doing all the work of designing full-fledged EU preferences. There is a field of robust decision-making (see, for example, Lempert 2014; Lempert et al. 2013) that explores the building of consensus on probabilities and preferences and to which welfare theory could perhaps usefully connect.

ACKNOWLEDGMENTS

This chapter is an abridged and edited version of Fleurbaey (2018), and both have benefitted from reactions of the audience at the 2016 Canadian Economic Association meeting in Ottawa and comments on the lecture or the paper by G. Dionne, M. Hoy, G. Ponthière, and a referee. Discussions with M. Ferranna have further improved this chapter.

REFERENCES

- Adler, Matthew D. 2013. *Well-Being and Fair Distribution*. Oxford: Oxford University Press.
- Adler, Matthew D., and Chris W. Sanchirico. 2006. "Inequality and uncertainty: Theory and legal applications." *University of Pennsylvania Law Review* 155: 279–377.
- Al-Najjar, Nabil I., and Jonathan Weinstein. 2009a. "The Ambiguity Aversion Literature: A Critical Assessment." *Economics and Philosophy* 25: 249–84.
- Al-Najjar, Nabil I., and Jonathan Weinstein. 2009b. "Rejoinder: The ambiguity aversion literature: A critical assessment." *Economics and Philosophy* 25: 357–69.

- Amiel, Yoram, Frank A. Cowell, and Wulf Gaertner. 2009. "To Be or Not to Be Involved: A Questionnaire-Experimental View on Harsanyi's Utilitarian Ethics." *Social Choice and Welfare* 32: 299–316.
- Arrow, Kenneth J. 1970. *Essays in the Theory of Risk-Bearing*. Amsterdam: North-Holland.
- Asheim, Geir B., and Stéphane Zuber. 2014. "Escaping the Repugnant Conclusion : Rank-Discounted Utilitarianism with Variable Population." *Theoretical Economics* 9: 629–50.
- Asheim, Geir B., and Stéphane Zuber. 2016. "Evaluating Intergenerational Risks." CES Working Paper, No. 2016.11.
- Blackorby, Charles, Walter Bossert, and David Donaldson. 1996. "Intertemporally consistent population ethics: Classical utilitarian principles." In Kenneth J. Arrow, Amartya K. Sen and Kotaro Suzumura (eds), *Social Choice Re-examined*, vol. 2 (London: Macmillan), pp. 137–62.
- Blackorby, Charles, David Donaldson, and Philippe Mongin. 2004. "Social Aggregation without the Expected Utility Hypothesis." Cahier # 2004-020, Laboratoire d'Econométrie, Ecole Polytechnique. Available at <https://hal.archives-ouvertes.fr/hal-00242932/document>.
- Bommier, Antoine, and Stéphane Zuber. 2008. "Can Preferences for Catastrophe Avoidance Reconcile Social Discounting with Intergenerational Equity?" *Social Choice and Welfare* 31: 415–34.
- Bradley, Richard. 2016. "Ellsberg's Paradox and the Value of Chances." *Economics and Philosophy* 32: 231–48.
- Broome, John. 1991. *Weighing Goods: Equality, Uncertainty and Time*. Oxford: Blackwell.
- Chambers, Christopher P., and Takashi Hayashi. 2014. "Preference Aggregation with Incomplete Information." *Econometrica* 82: 589–99.
- Coulhon, T., and P. Mongin 1989. "Social Choice Theory in the Case of von Neumann–Morgenstern Utilities." *Social Choice and Welfare* 6: 175–87.
- Debreu, Gérard. 1959. *Theory of Value: An Axiomatic Analysis of Economic Equilibrium*. New York: Wiley.
- Diamond, Peter A. 1967. "Cardinal Welfare, Individualistic Ethics, and Interpersonal Comparisons of Utility: Comment." *Journal of Political Economy* 75: 765–6.
- Durnin, Maureen, Michael Hoy, and Michael Ruse. 2012. "Genetic Testing and Insurance: The Complexity of Adverse Selection." *Ethical Perspectives* 19: 123–54.
- Dworkin, Ronald. 2000. *Sovereign Virtue: The Theory and Practice of Equality*. Cambridge, MA: Harvard University Press.
- Ellsberg, Daniel. 1961. "Risk, Ambiguity, and the Savage Axioms." *Quarterly Journal of Economics* 75: 643–69.
- Epstein, Larry G., and Uzi Segal. 1992. "Quadratic Social Welfare Functions." *Journal of Political Economy* 100: 691–712.
- Ferranna, Maddalena. 2015a. "Fairness in Cost Benefit Analysis: Equity-Enhanced Mean Variance Rules." Working paper, Faculty of Economics and Business, University of Leuven.
- Ferranna, Maddalena. 2015b. "Fairness, Risk and the Social Cost of Carbon." Working paper, Faculty of Economics and Business, University of Leuven.
- Ferranna, Maddalena. 2017. "Intergenerational Equity, Risk Aversion, and the Social Cost of Carbon." Mimeo.
- Finkelstein, Amy, Erzo F.P. Luttmer, and Matthew J. Notowidigdo. 2013. "What Good Is Wealth without Health? The Effect of Health on the Marginal Utility of Consumption." *Journal of the European Economic Association* 11: 221–58.
- Fleurbaey, Marc. 2008. *Fairness, Responsibility and Welfare*. Oxford: Oxford University Press.

- Fleurbaey, Marc. 2009. "Two Variants of Harsanyi's Aggregation Theorem." *Economics Letters* 105: 300–2.
- Fleurbaey, Marc. 2010. "Assessing Risky Social Situations." *Journal of Political Economy* 118: 649–80.
- Fleurbaey, Marc. 2016. "Rationality under Risk and Uncertainty." Mimeo.
- Fleurbaey, Marc. 2018. "Welfare Economics, Risk and Uncertainty." *Canadian Journal of Economics* 51: 5–40.
- Fleurbaey, Marc, Thibault Gajdos, and Stéphane Zuber. 2015. "Social Rationality, Separability and Equity under Uncertainty." *Mathematical Social Sciences* 73: 13–22.
- Fleurbaey, Marc, Marie Louise Leroux, and Gregory Ponthière. 2014. "Compensating the Dead." *Journal of Mathematical Economics* 51: 28–41.
- Fleurbaey, Marc, and François Maniquet. 2011. *A Theory of Fairness and Social Welfare*. Cambridge: Cambridge University Press.
- Fleurbaey, Marc, and Philippe Mongin. 2016. "The Utilitarian Relevance of the Aggregation Theorem." *American Economic Journal: Microeconomics* 8: 289–306.
- Fleurbaey, Marc, and Stéphane Zuber. 2015. "Discounting, Beyond Utilitarianism." *Economics: The Open-Access, Open-Assessment E-Journal* 9: 1–52. Available at <http://www.economics-ejournal.org/economics/discussionpapers/2014-40>.
- Fleurbaey, Marc, and Stéphane Zuber. 2017. "Fair management of social risk." *Journal of Economic Theory* 169: 666–706.
- Gajdos, Thibault, and Ferial Kandil. 2008. "The Ignorant Social Observer." *Social Choice and Welfare* 31: 193–232.
- Gajdos, Thibault, and Eric Maurin. 2004. "Unequal Uncertainties and Uncertain Inequalities: An Axiomatic Approach." *Journal of Economic Theory* 116: 93–118.
- Gajdos, Thibault, Jean-Marc Tallon, and Jean-Christophe Vergnaud. 2008. "Representation and Aggregation of Preferences under Uncertainty." *Journal of Economic Theory* 141: 68–99.
- Gilboa, Itzhak, Andrew Postlewaite, and David Schmeidler. 2009. "Is It Always Rational to Satisfy Savage's Axioms?" *Economics and Philosophy* 25: 285–96.
- Gilboa, Itzhak, and David Schmeidler. 1989. "Maxmin Expected Utility with a Non-Unique Prior." *Journal of Mathematical Economics* 18: 141–53.
- Gollier, Christian, Bruno Jullien, and Nicolas Treich. 2000. "Scientific Progress and Irreversibility: An Economic Interpretation of the 'Precautionary Principle.'" *Journal of Public Economics* 75: 229–253.
- Grant, Simon, Atsushi Kajii, Ben Polak, and Zvi Safra. 2010. "Generalized Utilitarianism and Harsanyi's Impartial Observer Theorem." *Econometrica* 78: 1939–71.
- Grant, Simon, Atsushi Kajii, Ben Polak, and Zvi Safra. 2011. "Equally Distributed Equivalent Utility, *Ex post* Egalitarianism and Utilitarianism." *Journal of Economic Theory* 147: 1545–71.
- Grant, Simon, Atsushi Kajii, Ben Polak, and Zvi Safra. 2012. "A Generalized Representation Theorem for Harsanyi's ('Impartial') Observer." *Social Choice and Welfare* 39: 833–46.
- Hammond, Peter J. 1982. "Utilitarianism, Uncertainty and Information." In Amartya K. Sen and Bernard Williams (eds), *Utilitarianism and Beyond* (Cambridge: Cambridge University Press), pp. 85–102.
- Hammond, Peter J. 1987. "On Reconciling Arrow's Theory of Social Choice with Harsanyi's Fundamental Utilitarianism." In George R. Feiwel (ed.), *Arrow and the Foundations of the Theory of Economic Policy* (New York: New York University Press), pp. 179–221.
- Harsanyi, John C. 1953. "Cardinal Utility in Welfare Economics and in the Theory of Risk Taking." *Journal of Political Economy* 61: 434–5.

- Harsanyi, John C. 1955. "Cardinal Welfare, Individualistic Ethics, and Interpersonal Comparisons of Utility." *Journal of Political Economy* 63: 309–21.
- Harsanyi, John C. 1977. *Rational Behavior and Bargaining Equilibrium in Games and Social Situations*. Cambridge: Cambridge University Press.
- Heal, Geoffrey, and Antony Millner. 2015. "Should Climate Policy Account for Ambiguity?" LSE Grantham Institute Working Paper 202.
- Herne, Kaisa, and Maria Suojanen. 2004. "The Role of Information in Choices over Income Distributions." *Journal of Conflict Resolution* 48: 173–93.
- Hild, Matthias, Richard Jeffrey, and Mathias Risse. 2008. "Preference Aggregation after Harsanyi." In Marc Fleurbaey, Maurice Salles, and John A. Weymark (eds), *Justice, Political Liberalism, and Utilitarianism* (Cambridge: Cambridge University Press), pp. 198–218.
- Hoy, Michael. 2006. "Risk Classification and Social Welfare." *Geneva Papers on Risk and Insurance* 31: 245–269.
- Karni, Edi. 1998. "Impartiality: Definition and Representation." *Econometrica* 66: 1405–15.
- Karni, Edi, and John A. Weymark. 1998. "An Informationally Parsimonious Impartial Observer Theorem." *Social Choice and Welfare* 15: 321–32.
- Keeney, Ralph L. 1980. "Equity and Public Risk." *Operations Research* 28: 527–34.
- Kocher, Martin G., Amrei M. Lahno, and Stefan T. Trautmann. 2015. "Ambiguity Aversion Is the Exception." CESifo Working Paper Series Number 5261.
- Kolm, Serge-Christophe. 1996. *Modern Theories of Justice*. Cambridge, MA: MIT Press.
- Leach, John. 2010. "Ex post Welfare under Alternative Health Care Systems." *Journal of Public Economic Theory* 12: 1027–57.
- Lempert, Robert J. 2014. "Embedding (Some) Benefit-Cost Concepts into Decision Support Processes with Deep Uncertainty." *Journal of Benefit-Cost Analysis* 5: 487–514.
- Lempert, Robert J., David G. Groves, and Jordan R. Fischbach. 2013. "Is It Ethical to Use a Single Probability Density Function?" Santa Monica, CA: RAND Corporation. https://www.rand.org/pubs/working_papers/WR992.html.
- Machina, Mark J. 1989. "Dynamic Consistency and Non-Expected Utility Models of Choice under Uncertainty." *Journal of Economic Literature* 27: 1622–68.
- Mirrlees, J.A. 1982. "The economic uses of utilitarianism." In Amartya K. Sen and Bernard Williams (eds), *Utilitarianism and Beyond* (Cambridge: Cambridge University Press), pp. 63–84.
- Mongin, Philippe. 2001. "The Impartial Observer Theorem of Social Ethics." *Economics and Philosophy* 17: 147–80.
- Mongin, Philippe, and Claude d'Aspremont. 1998. "Utility Theory and Ethics." In Salvador Barbera, Peter Hammond, and Christian Seidl (eds), *Handbook of Utility Theory, Volume 1: Principles* (Dordrecht: Kluwer), pp. 371–481.
- Mongin, Philippe, and Marcus Pivato. 2015. "Ranking Multidimensional Alternatives and Uncertain Prospects." *Journal of Economic Theory* 157: 146–71.
- Mongin, Philippe, and Marcus Pivato. 2016. "Social Evaluation under Risk and Uncertainty." In Matthew D. Adler and Marc Fleurbaey (eds), *The Oxford Handbook of Well-Being and Public Policy* (Oxford: Oxford University Press), pp. 711–45.
- Parfit, Derek. 1991. "Equality or Priority?" The Lindley Lecture, University of Kansas. <https://kuscholarworks.ku.edu/handle/1808/12405>.
- Pattanaik, Prasanta K. 1968. "Risk, Impersonality, and the Social Welfare Function." *Journal of Political Economy* 76: 1152–69.
- Rawls, John. 1971. *A Theory of Justice*. Cambridge, MA: Harvard University Press.
- Rheinberger, Christophe M., and Nicolas Treich. 2015. "Attitudes Toward Catastrophes." Working Paper TSE 635.

- Roemer, John E. 1996. *Theories of Distributive Justice*. Cambridge, MA: Harvard University Press.
- Roemer, John E. 2006. "The Global Welfare Economics of Immigration." *Social Choice and Welfare* 27: 311–25.
- Savage, Leonard J. 1972. *The Foundations of Statistics*. 2nd ed. New York: Dover.
- Sen, Amartya K. 1974. "Rawls versus Bentham: An Axiomatic Examination of the Pure Distribution Problem." *Theory and Decision* 4: 301–9.
- Sen, Amartya K. 1977. "Non-Linear Social Welfare Functions: A Reply to Professor Harsanyi." In Robert E. Butts and Jaakko Hintikka (eds), *Foundational Problems in the Special Sciences* (Dordrecht: Reidel), pp. 297–302.
- Thomson, William. 2011. "Fair Allocation Rules." In Kenneth J. Arrow, Amartya K. Sen, and Kotaro Suzumura (eds), *Handbook of Social Choice and Welfare*, vol. 2 (Amsterdam: North-Holland), pp. 393–506.
- Traub, Stefan, Christian Seidl, Ulrich Schmidt, and Maria V. Levati. 2005. "Friedman, Harsanyi, Rawls, Boulding – or Somebody Else? An Experimental Investigation of Distributive Justice." *Social Choice and Welfare* 24: 283–309.
- Voorhoeve, Alex, Ken Binmore, Arnaldur Stefansson, and Lisa Stewart. 2016. "Ambiguity Attitudes, Framing, and Consistency." *Theory And Decision* 81: 313–37.
- Voorhoeve, Alex, and Marc Fleurbaey. 2016. "Priority or Equality for Possible People?" *Ethics* 126: 929–54.
- Weymark, John A. 1991. "A Reconsideration of the Harsanyi-Sen Debate on Utilitarianism." In Jon Elster and John Roemer (eds), *Interpersonal Comparisons of Well-Being* (Cambridge: Cambridge University Press), pp. 255–320.
- Zuber, Stéphane. 2016. "Harsanyi's Theorem without the Sure-Thing Principle: On the Consistent Aggregation of Monotonic Bernoullian and Archimedean preferences." *Journal of Mathematical Economics* 63: 78–83.

CHAPTER 21

THE ETHICS OF MAKING RISKY DECISIONS FOR OTHERS

LUC BOVENS

21.1 INTRODUCTION

UTILITARIANISM, it is often said, is not sensitive to the distribution of welfare across different individuals.¹ In a risk-free environment, what this means is that utilitarianism is indifferent whether an increment in welfare of a given size goes to a better-off or a worse-off person, and it requires an intervention that procures even a slightly greater increase in welfare to the better off at the cost of a smaller decrease in welfare to the worse off.

One can construe this as a kind of dilemma. On the one hand, the utilitarian has a reasonable objective, that is, to maximize total or average welfare in society. On the other hand, the sensitivity of their opponent is reasonable as well: the worse off should not be expected to incur losses for the sake of just marginally larger gains to the better off. Different decision-makers will resolve this dilemma in different ways depending on how sensitive they are to the distribution of welfare.

In the presence of risk, the matter is much more complex than in a risk-free environment because there are more distributions in play that elicit distinct sensitivities. These sensitivities will generate different types of dilemmas.

In Section 21.2, I will present four dilemmas that all involve risk in decision-making for others. These dilemmas seem quite disparate. In Section 21.3, I construct a model that captures four distinct distributional sensitivities that a decision-maker might have when making risky decisions for others. The model specifies four parameters that quantify the

¹ For example, see Gauthier (1963: 121–7), Rawls (1974–1975), Sen (1973), and Williams (1973).

strength of these sensitivities. In Section 21.4, I will show how each of the four dilemmas occurs because a particular distributional sensitivity kicks in and pulls us away from what the utilitarian would prescribe in the context of decision-making under risk. In Section 21.5, I will turn to the much-discussed case of *Saving the Few at Greater Risk to the Many*. This case is more complex, because there are two distributional sensitivities at work that pull in opposite directions from the utilitarian calculus. Finally, in Section 21.6, I will consider objections to each of the four distributional sensitivities. Section 21.7 reflects on the role of formal modeling by means of prospects, and Section 21.8 concludes.

21.2 FOUR DILEMMAS

21.2.1 Restrictive Intervention

Suppose that we have a cheap but less effective drug that we can provide to all, or an expensive but more effective drug that, due to budgetary constraints, we can provide only to half the population. The catch is that the expected number of lives saved by the more expensive drug program, even if only provided to half, is greater than the expected number of lives saved by the cheaper drug program, provided to all. Should we procure the cheaper drug to all or the more expensive drug to half?

Ubel et al. (1996) construct a case in which the allocation of the expensive drug is conducted by randomization. They polled decision theorists, prospective jurors, and medical ethicists asking whether we should procure the cheaper drug to all or the more expensive drug to half. This case constitutes a dilemma with forces pulling in opposite directions. On the one hand, one wants to maximize the expected number of people saved. This favors the more expensive drug. On the other hand, once the randomization device has determined who will get the more expensive drug, people have unequal chances of survival. This favors the cheaper drug. Indeed, Ubel et al. find that decision theorists tend to favor the more expensive drug for half, while prospective jurors and medical ethicists tend to favor the cheaper drug for all.

21.2.2 Diversification

The *GiveWell* website (<http://www.givewell.org>), run by Effective Altruists, recommends charities that provide the greatest expected reduction in suffering and premature death for each marginal dollar contributed. Organizations that do extremely well, according to this calculus, are anti-malaria and deworming initiatives.

Critics ask why *GiveWell's* calculus is not more favorable to otherwise highly regarded charities like Oxfam, MSF (*Médecins Sans Frontières*), and the Red Cross. The response from Effective Altruists is that they respect these organizations, but that organizations

that support diverse causes do worse in their calculus than organizations that support single causes. One reason for this is that a diversified portfolio tends to include causes with effects that are hard to measure, such as advocacy against corruption.

However, there is another difference between single-cause and multiple-cause organizations. For single-cause organizations, the chances that your marginal dollar will do some good are more correlated than for multiple-cause organizations.

Consider deworming initiatives. These initiatives support massive deworming in schools in the developing world. The pupils are not tested individually, because treatment is cheaper than testing and the pills are relatively safe. The original study by Miguel and Kremer (2003) suggested that deworming has huge educational impacts (which hold the promise of future economic benefits). Taylor-Robinson et al. (2015), Aiken et al. (2015), and Davey et al. (2015) cast doubt on these results. However, *GiveWell* continues to support the initiatives, because they are cost effective. Deworming is cheap and it just *might* have a huge impact. However, a blogger on the *GiveWell* website also admits that it “might have close to zero impact” (Sean 2017).

Granted, deworming may not work equally well in all places, but one would expect that the success or failure of these interventions around the developing world is at least to some extent correlated. Hence, in giving to deworming initiatives, there is a non-negligible chance that one’s donations will have little or no impact. By contrast, contributions to organizations like Oxfam go to a broad array of causes. The expected reduction in suffering and premature death from donations may be lower than from deworming initiatives, but at least a portion of one’s donation is likely to do some good.

The dilemma is this: should we make charitable donations to single-cause organizations in order to maximize the expected benefits, even if there is a good chance that our donations will be money down the drain? Or, should we donate to multiple-cause organizations in order to make sure that our donations will at least do some good in the world, even if the expected benefits of our donations are lower?

21.2.3 Unwelcome Risk Reductions

Spiegelhalter (2015) has constructed a flow graph that offers a visual representation of the costs and benefits of breast cancer screening in the UK. The graph shows that 7.5 percent of women between 50 and 70 in the UK have symptoms of breast cancer that would be discovered by screening. If they are screened, then all 7.5 percent will be treated, 6 percent will survive and 1.5 percent will die. If they are not screened, then 1.5 percent will have symptoms that would be discovered in screening, but they will never become aware of it: they will remain undiagnosed, untreated, and the disease will never bother them. In other words, if they had been screened, this 1.5 percent would have been needlessly treated. The remaining 6 percent will eventually show symptoms and be treated; 4 percent will survive and 2 percent will die.

So, screening reduces the chance of mortality with $2\text{ percent} - 1.5\text{ percent} = 0.5\text{ percent}$. But it comes at a cost of 1.5 percent who are needlessly treated. If that 1.5 percent would

not have been screened, they would have never known that they had symptoms of breast cancer.

Many women opt against screening when told about these statistics. A mere 0.5 percent reduction in mortality does not justify the hassle for them. There is the annoyance of the screening, the needless agony of false positives, and the 1.5 percent of treatments that are unnecessary. And yet, screening does save about 1300 lives per year in the UK.

This is our dilemma. On the one hand, it seems that we should respect the wishes of fully informed patients. On the other hand, there is a compelling case to be made that, taking into account costs, saving an expected 1300 lives merits the implementation of a screening process.

21.2.4 Not on My Watch

Johnson and Rehavi (2016) present a study of the percentage of deliveries by Caesarean sections (or C-sections) for prospective mothers who are themselves MDs versus for prospective mothers who are not MDs. They find that the percentage of unscheduled C-sections—that is after a vaginal birth was attempted—is higher for prospective mothers who are not MDs.

Why is this the case? It may be the case that some obstetricians need to make their car payments by the end of the month and perform needless C-sections. C-sections are more lucrative than vaginal births and under conditions of asymmetry of information, financial incentives create physician induced demand for C-sections. However, they cannot pull this off on prospective mothers who are MDs, because those patients know better. One datum in support of this hypothesis is that we also find a lower rate of C-sections on partners of prospective fathers who are MDs, and that the rate of C-sections is even lower when we focus on prospective mothers who are obstetricians. The results do not hold for scheduled C-sections, that is, C-sections that were planned beforehand. The authors suggest that this is the case because there are strict criteria for scheduled C-sections.

Another explanation is that the obstetricians are afraid of being sued by the women or their families for *not* proceeding with a C-section during somewhat difficult or risky deliveries. They may trust that prospective mothers or partners of prospective fathers who are themselves MDs are less likely to file suit against a fellow doctor for medical malpractice.

However, let us try to be less cynical and attempt a friendlier explanation. When there are concerns during the birthing process, proceeding to a C-section is often the less risky option. Vaginal birth can result in success (a healthy mother and foetus) or failure (death or injury to the mother or foetus). Due to the longer recovery period, the outcome of a C-section typically ranks between an outcome of success and an outcome of failure of a vaginal birth.

If the prospective mother is not an MD, then the obstetrician takes the decision *for* the prospective mother. Even if holding out for a vaginal birth is a risk that the obstetrician

themselves would want to take if they were in the position of the prospective mother, they may not want to take this risk *for* the prospective mother. Hence, they proceed to a C-section more swiftly.

To defend this decision, they may say: “I can’t have death or injury due to forgoing a C-section happen *on my watch*.” And this is not (just) because they are concerned about their reputation or about being sued. Rather, a good person finds it harder to take a loss when gambling with the welfare of others than with the welfare of themselves. This is why they might want to take the gamble when watching over themselves, but not when watching over others.

If the prospective mother is herself an MD, then the obstetrician is more like a fiduciary care-taker. They simply adopt the risk attitudes of the MD prospective mother. The obstetrician does not take the decision *for* her—rather they see themselves as executing the decision that the prospective mother would take for herself. Hence, they are less swift in proceeding to a C-section.²

The dilemma can then be phrased as follows. The presumption is that non-MD prospective mothers, if they were fully informed, would have risk attitudes that are no different from MD prospective mothers. Should an obstetrician respect the risk attitudes that the non-MD prospective mother would have if she were fully informed? Or, is it permissible for the obstetrician to import additional risk aversion over and above the risk attitudes of the non-MD prospective mother?

21.3 DISTRIBUTIONAL SENSITIVITIES

I will construct a model to cast light on where the pull comes from in each of these dilemmas.

21.3.1 Four Distributional Sensitivities

Suppose that you were asked to make a choice between prospects *Risky* and *Certain* in Tables 21.1 and 21.2 that will affect two people. A matrix represents a *prospect*, rows represent *personal prospects*, and columns represent *states* (of the world). There is a probability distribution defined over the states and the entries in the matrix are utility values.

We can calculate the *value* of the prospect *ex ante*, that is, by computing the average of the expected utilities. Or, we can calculate the value of the prospect *ex post*, that is, by computing the expectation of the average utility in each state. By simple algebra, the average of the expectations equals the expectation of the state averages.

² Granted, one problem with this explanation is that the difference in C-section rates performed on non-MDs versus MDs disappears when we restrict ourselves to HMO hospitals in which obstetricians get a fixed salary and do not work on a fee-for-service basis. This favors the hypothesis of physician-induced demand driven by financial incentives.

Table 21.1 Risky

Expected Utility (EU)		S1 <i>Prob(S1) = .3</i>	S2 <i>Prob(S1) = .7</i>
6.7	Person 1	20	1
3.4	Person 2	2	4
Average EU	State Average (SA)	11	2.5
Expectation of SA			
5.05			

Table 21.2 Certain

Expected Utility		S <i>Prob(S) = 1</i>
5	Person 1	5
5	Person 2	5
Average EU	State Average (SA)	5
Expectation of SA		
5		

Harsanyi (1955) imposes a set of requirements of minimal rationality on the ranking of prospects and shows, in his aggregation theorem, that we should rank prospects relative to the *weighted* average of the expectations in the prospect. If we add to this a requirement of anonymity—that is, that the identity of the people does not matter—then we are required to order prospects relative to the average of the expectations of the persons in the prospect. Hence, following Harsanyi, we should rank *Risky* over *Certain*.

But forget about theory for a second. Suppose that a decision-maker is entrusted with making a decision between these two prospects. I submit that many decision-makers would want to play it safe and choose *Certain* over *Risky*. What might such a decision-maker say if they were asked to give a justification for their decision? They might mention a number of things that they dislike about *Risky*:

- (i) They may dislike the fact that person 2's expectation is substantially lower than person 1's.
- (ii) They may dislike the fact that state S2, which is likely to materialize, has a low average utility.
- (iii) They may dislike the fact that whatever state materializes, there will be inequality between the persons.
- (iv) They may dislike taking the risk of receiving a low outcome on behalf of person 1 if S2 materializes or a low outcome on behalf of person 2 if S1 materializes.

The decision-maker finds fault with four types of distributions: in (i), they find fault with the *inter-personal prospect distribution*. This is the distribution of the expected utilities, viz. $\langle 6.7, 3.4 \rangle$. In (ii), they find fault with the *inter-state distribution*. This is the distribution of the state averages, viz. $\langle 11, 2.5 \rangle$. In (iii), they find fault with the *intra-state distributions*. There are two intra-state distributions, viz. $\langle 20, 2 \rangle$ and $\langle 1, 4 \rangle$. In (iv), they find fault with the *intra-personal prospect distributions*. There are two intra-personal prospect distributions, viz. $\langle 20, 1 \rangle$ and $\langle 2, 4 \rangle$.

The decision-maker resists the recommendation of the prospect with the greatest average expectation or the greatest expectation of averages by pointing to four worrisome distributional features. They are sensitive to spread, that is, they would rather see less than more spread in these distributions and this is what motivates the decision-maker to recommend B over A. For simplicity, let us name the average of the expectations or the expectation of the state averages the *utilitarian* value of the prospect, though the import of Harsanyi's Theorem is open to interpretation (see Weymark 1991 and Greaves 2017). Hence, the utilitarian value of *Risky* may exceed the utilitarian value of *Certain* and yet, we would prefer *Certain* over *Risky* on distributional grounds.

21.3.2 Modeling Distributional Sensitivities

I will import these distributional sensitivities in the determination of the value of a prospect.

21.3.2.1 *The Inter-Personal Prospect Distribution*

Suppose that we are sensitive to the inter-personal prospect distribution. Let us take some simple numbers. Suppose that person 1 has an expectation of 16 and person 2 has an expectation of 4, so that the utilitarian value of the prospect is 10.

A decision-maker who would rather see less of a spread in the distribution of expected utility might say: "This prospect has the same value to me as a prospect in which each were to have the same expectation—an expectation that is lower than 10 and greater than or equal to 4." If they are slightly sensitive to the spread in the distribution, then this expectation will be a number closer to 10. If they are more sensitive, then this expectation will be a number closer to 4.

What we need is a function with a single parameter α so that, as α moves from 0 to $+\alpha$, the value of the distribution moves from 10 to 4. The following function does precisely this.³ For $i = 1, \dots, n$ persons, let x_i be the expectation of person i and, to respect anonymity, let us set the person-weights $w_i = 1/n$.

³ The function is in Atkinson (1970: 249–52) and is applied to prospects in Fleurbaey (2010: 658). This function is not rank-order sensitive. Alternatively, we can use a rank-order sensitive function as described in Donaldson and Weymark (1980: 74). See Bovens (2015a).

$$\begin{aligned} & \left(\sum_{i=1}^n w_i x_i^{(1-\alpha)} \right)^{1/(1-\alpha)} \quad \text{for } \alpha \neq 1 \\ & \exp \left(\sum_{i=1}^n w_i \ln x_i \right) \quad \text{for } \alpha = 1 \end{aligned} \quad (1)$$

This function yields the *equally distributed equivalent* of the distribution.

To illustrate, set α at the intermediate value of $1/2$. Then in our example,

$$\left(1/2\sqrt{16} + 1/2\sqrt{4} \right)^2 = 9 \quad (2)$$

Because we apply the function to expectations here, it yields the *equally-distributed equivalent of the expectations* (EDEE) for a decision-maker with the inter-personal prospect distribution-sensitivity α_{EDEE} . So, we could say that the inter-personal prospect distribution-sensitive value of the prospect is not 10, but rather 9 for a decision-maker characterized by the sensitivity parameter $\alpha_{EDEE} = 1/2$.

Returning to our prospects *Risky* and *Certain* in Tables 21.1 and 21.2, the utilitarian value of the prospects is the average of the expectations, or, in other words, the equally distributed equivalent of the expectations with $\alpha_{EDEE} = 0$. Hence *Risky* ranks higher than *Certain* on the utilitarian calculus. However, the inter-personal prospect distribution-sensitive value with $\alpha_{EDEE} = 1/2$ of the prospects is:

$$V_{\alpha_{EDEE}=1/2}(\text{Risky}) = \left(1/2\sqrt{6.7} + 1/2\sqrt{3.4} \right)^2 = 4.91 \dots < 5 = V_{\alpha_{EDEE}=1/2}(\text{Certain}) \quad (3)$$

Hence, a decision-maker with this type of sensitivity would choose *Certain* over *Risky*.

This is one distributional sensitivity. The decision-maker may bring other distributional sensitivities to the evaluation of the prospects. We will treat each of these distributional sensitivities in the same formal manner.

21.3.2.2 The Inter-State Distribution

Consider the inter-state distribution sensitivity. Suppose that there are two equiprobable states with average utilities 16 and 4. Then we could do the same exercise. The utilitarian is not sensitive to this distribution and the value of the prospect is just the expectation of the state averages, viz. 10. A decision-maker who is slightly sensitive to the inter-state distribution would consider the value of the prospect to be lower than but closer to 10, whereas a decision-maker who is very sensitive to the inter-state distribution would consider the value of the prospect to be closer to 4. The following function permits us to model inter-state sensitivities with x_j being state averages for states $j = 1, \dots, m$ and β ranging from 0 to $+\infty$.

$$\begin{aligned} & \left(\sum_{j=1}^m p_j x_j^{(1-\beta)} \right)^{1/(1-\beta)} \quad \text{for } \beta \neq 1 \\ & \exp \left(\sum_{j=1}^m p_j \ln x_j \right) \quad \text{for } \beta = 1 \end{aligned} \quad (4)$$

This function yields the *certainty equivalent* of the distribution.

For β equals $\frac{1}{2}$, the value of the prospect with averages 16 and 4 in equiprobable states equals 9. Because we apply the function to state averages, it yields the *certainty equivalent of the averages* (CEA) for a person with inter-state sensitivity $\beta_{CEA} = \frac{1}{2}$.

Returning to our prospects *Risky* and *Certain*, the inter-state distribution-sensitive value of the prospects is:

$$V_{\beta_{CEA}=1/2}(\text{Risky}) = (.3\sqrt{11} + .7\sqrt{2.5})^2 = 4.42\dots < 5 = V_{\beta_{CEA}=1/2}(\text{Certain}) \quad (5)$$

So, the decision-maker with this type of sensitivity would choose *Certain* over *Risky*.

21.3.2.3 The Intra-State Distribution

Suppose that the decision-maker is sensitive to the intra-state distribution. In this case, they will calculate the *equally-distributed equivalent* (EDE) of each state and set the value of the prospect at the expectation of the EDEs. Again, let us set $\alpha_{EDE} = 1/2$. In our example:

$$V_{\alpha_{EDE}=1/2}(\text{Risky}) = .3\left(1/2\sqrt{20} + 1/2\sqrt{2}\right)^2 + .7\left(1/2\sqrt{1} + 1/2\sqrt{4}\right)^2 = 4.17 \quad (6)$$

which is smaller than $V_{\alpha_{EDE}=1/2}(\text{Certain}) = 5$.

21.3.2.4 The Intra-Personal Prospect Distribution

Finally, suppose that the decision-maker is sensitive to the intra-personal prospect distribution. Then they will calculate the *certainty-equivalents* (CE) for each person and set the value of the prospect at the average of the CEs. In our example with $\beta_{CE} = 1/2$:

$$V_{\beta_{CE}=1/2}(\text{Certain}) = 1/2\left(.3\sqrt{20} + .7\sqrt{1}\right)^2 + 1/2\left(.3\sqrt{2} + .7\sqrt{4}\right)^2 = 3.75 \quad (7)$$

which is smaller than $V_{\beta_{CE}=1/2}(\text{Risky}) = 5$.

Table 21.3 provides an overview.

Table 21.3 Distributional Sensitivities, Prospect Values, and Parameters

Distributional Sensitivity	Value of the Prospect	Parameter
Inter-Personal Prospect	Equally Distributed Equivalent of the Personal Prospect Expectations	α_{EDE}
Inter-State	Certainty Equivalent of the State Averages	β_{CEA}
Intra-State	Expectation of the Equally Distributed Equivalents of the States	α_{EDE}
Intra-Personal Prospect	Average of the Certainty Equivalents of the Personal Prospects	β_{CE}

21.4 REVISITING THE DILEMMAS

I will now revisit each of the dilemmas of Section 21.2. I will construct a stylized model for each dilemma and show how the distributional sensitivities underlie these dilemmas.

21.4.1 Sensitivity to the Inter-Personal Prospect Distribution and Restrictive Intervention

We model the case in which the people have been allocated to the groups that will and will not receive the more expensive drug in Table 21.4. Suppose that without any intervention, there is a 0.70 survival rate. A cheaper intervention provided to all can increase the survival rate to 0.80. A more expensive intervention provided to half the population can increase the survival rate to $0.90 + \varepsilon$. Let us first model the allocation without randomization. For example, we might stipulate that the more expensive drug is provided to the urban population, but not to the equally large rural population.

A utilitarian decision-maker prefers *Restrictive Intervention* to *Intervention for All*.

$$V_{UTIL}(RestrictiveIntervention) = 0.80 + \varepsilon/2 > 0.80 = V_{UTIL}(InterventionAll) \quad (8)$$

But now suppose that the decision-maker is sensitive to the inter-personal prospect distribution. For sufficiently large values of α_{EDEE} and sufficiently small values of ε , the value *Intervention for All* exceeds the value of *Restrictive Intervention* and hence they will prefer *Intervention for All*.

$$V_{\alpha_{EDEE}=x>0}(RestrictiveIntervention) < V_{\alpha_{EDEE}=x>0}(InterventionAll) \quad (9)$$

In Ubel et al. (1996), the allocation is done by randomization. The argument above still applies if we focus on the time-point after the randomization has taken place. One

Table 21.4 *Restrictive Intervention and Intervention for All*

<table><tr><td>EU</td><td></td><td></td><td></td><td></td><td></td></tr><tr><td>0.90</td><td>0</td><td>1</td><td>1</td><td>1</td><td>...</td></tr><tr><td>+ ε</td><td></td><td></td><td></td><td></td><td></td></tr><tr><td>...</td><td>...</td><td>...</td><td>...</td><td>...</td><td>...</td></tr><tr><td>0.70</td><td>0</td><td>0</td><td>0</td><td>1</td><td>...</td></tr><tr><td>...</td><td>...</td><td>...</td><td>...</td><td>...</td><td>...</td></tr></table>	EU						0.90	0	1	1	1	...	+ ε						...	...	...	...	...	...	0.70	0	0	0	1	...	...	...	...	...	...	...	$\succ_{UTIL}$ $\prec_{InterPP}$	<table><tr><td>EU</td><td></td><td></td><td></td><td></td><td></td></tr><tr><td>0.80</td><td>0</td><td>0</td><td>1</td><td>...</td><td></td></tr><tr><td>...</td><td>...</td><td>...</td><td>...</td><td>...</td><td>...</td></tr><tr><td>...</td><td>...</td><td>...</td><td>...</td><td>...</td><td>...</td></tr><tr><td>...</td><td>...</td><td>...</td><td>...</td><td>...</td><td>...</td></tr><tr><td>0.80</td><td>1</td><td>0</td><td>0</td><td>...</td><td></td></tr></table>	EU						0.80	0	0	1	...		...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	0.80	1	0	0	...	
EU																																																																										
0.90	0	1	1	1	...																																																																					
+ ε																																																																										
...	...	...	...	...	...																																																																					
0.70	0	0	0	1	...																																																																					
...	...	...	...	...	...																																																																					
EU																																																																										
0.80	0	0	1	...																																																																						
...	...	...	...	...	...																																																																					
...	...	...	...	...	...																																																																					
...	...	...	...	...	...																																																																					
0.80	1	0	0	...																																																																						
<i>Restrictive Intervention</i>		<i>Intervention for All</i>																																																																								

might say that the medical ethicists and the prospective jurors, who favor *Intervention for All*, focus on the post-randomization point when the persons in the prospect face unequal expectations in their personal prospects and are sensitive to the inter-personal prospect distribution. The decision theorists do not have this sensitivity and follow the utilitarian calculus.

The decision theorists could object that in case of randomization, *Restrictive Intervention* is really the only reasonable solution. If we focus on the time point before the randomization took place, each person's expectation from this program is both (i) equal and (ii) greater than each person's expectation from the program providing the cheaper drug to the whole population. Hence, the expensive drug should win out and sensitivity to the inter-personal prospect distribution cannot undo this.

The medical ethicists and prospective jurors could retort: Why should the preferred reference point be the pre-randomization point? The decision theorist could respond that participants *would have* consented if they had been asked before the randomization. But the fact of the matter is that they were not asked and why should *hypothetical* consent qualify as a justification? Someone who favors *Intervention for All* because they take the post-randomization inequality to be morally relevant is not making a mistake or is not being irrational.

21.4.2 Sensitivity to the Inter-State Distribution and Diversification

Imagine a population in desperate need of a charity intervention. In the model, the outcome of a failed intervention is assigned the value 0 (say, for death) and the outcome of a successful outcome is assigned the value 1 (say, for full health). *Give Well* recommends charities that are single cause and have a reasonable chance of success. The chance that the intervention is successful is, say, $2/3 + \epsilon$. If the intervention is successful, then it is successful for all. *Oxfam* supports multiple causes. These interventions offer slightly less of a chance of success to each person—say, a chance of $2/3$. But they are sufficiently varied so that they will not all fail together. *Multiple-Cause* in Table 21.5 is represented as

Table 21.5 <i>Single-Cause and Multiple-Cause</i>								
				$\succ_{UTIL}$ $\prec_{InterState}$				
EU	$P(S_1) = 1/3 - \epsilon$	$P(S_2) = 2/3 + \epsilon$			EU	$P(S_1) = 1/3$	$P(S_2) = 1/3$	$P(S_3) = 1/3$
$2/3 + \epsilon$	0	1	1		$2/3$	0	1	1
$2/3 + \epsilon$	0	1	1		$2/3$	1	0	1
$2/3 + \epsilon$	0	1	1		$2/3$	1	1	0
SA	0	1	1		SA	$2/3$	$2/3$	$2/3$
<i>Single-Cause</i>					<i>Multiple-Cause</i>			

anti-correlated risk. (Here and in the tables below I represent prospects with anti-correlated risk rather than with independent risk. I do so for simplicity of presentation, because the contrast I wish to draw holds equally well for anti-correlated risk as for independent risk.)

A utilitarian decision-maker prefers *Single-Cause* over *Multiple-Cause*:

$$V_{UTIL}(SingleCause) = 2/3 + \varepsilon > 2/3 = V_{UTIL}(MultipleCause) \quad (10)$$

The decision-maker who favors *Multiple-Cause* deviates from the utilitarian recommendation by adopting a sensitivity to the inter-state distribution. For a sufficiently large value of the parameter β_{CEA} and a sufficiently small ε ,

$$V_{\beta_{CEA}=x>0}(SingleCause) < V_{\beta_{CEA}=x>0}(MultipleCause) \quad (11)$$

They look at the state averages and are reluctant to support a single-cause charity that has a reasonable chance of being good for nothing. Deworming projects may have great value on the utilitarian calculus, but if the research on deworming does not bear out, as has been suggested in follow-up studies, then all the donations are money down the drain. The utilitarian value of single-cause charities may be greater, but the certainty equivalent of the state averages is greater for multiple-cause charities than for single-cause charities. In other words, they would rather play the short odds of Oxfam or MSF than the long odds of deworming initiatives, even though the expected value of the latter is greater than of the former.

21.4.3 Sensitivity to the Intra-State Distribution and Unwelcome Risk Reductions

In Table 21.6, we construct a stylized model of the breast cancer screening case that uses different numbers than in Spiegelhalter's flow graph, but still captures the dilemma associated with screening for breast cancer in public health. Suppose that without screenings, the chances of survival are $2/3$ and risks are independent. With screenings,

Table 21.6 *No Screening and Screening*

EU				$>_{UTIL}$ $<_{IntraState}$	EU			
$2/3$	0	1	1		$2/3 - \varepsilon$	$2/3 - \varepsilon$	$2/3 - \varepsilon$	$2/3 - \varepsilon$
$2/3$	1	0	1		$2/3 - \varepsilon$	$2/3 - \varepsilon$	$2/3 - \varepsilon$	$2/3 - \varepsilon$
$2/3$	1	1	0		$2/3 - \varepsilon$	$2/3 - \varepsilon$	$2/3 - \varepsilon$	$2/3 - \varepsilon$
No Screening					Screening			

there are no cancer fatalities. However, there are still costs to screenings: the mere imposition, the false positives, and the unnecessary surgeries. We set the utility of screening at $2/3 - \epsilon$, so that each woman would opt for *No Screening* over *Screening*.

A utilitarian decision-maker who is insensitive to the intra-state distribution prefers *No Screening* to *Screening*.

$$V_{UTIL}(NoScreening) > V_{UTIL}(Screening) \tag{12}$$

However, a decision-maker who is sensitive to the intra-state distribution dislikes the fact that, under *No Screening*, there is a wide spread in the utilities within each state. If the equally distributed equivalent of each state in *No Screening* drops below $2/3 - \epsilon$, then the value of the prospect *No Screening* will be lower than the value of the prospect *Screening*. This will indeed be the case for sufficiently large α_{EDE} and sufficiently small ϵ .

$$V_{\alpha_{EDE}=x>0}(NoScreening) < V_{\alpha_{EDE}=x>0}(Screening) \tag{13}$$

So, a decision-maker with a sensitivity for the intra-state distributions will overrule the preference of the women affected. To justify their decision, they will point to the fact that there is bound to be a large casualty rate within each state.⁴

21.4.4 Sensitivity to the Intra-Personal Prospect Distribution and Not on My Watch

Suppose that, given certain obstetric indications, an obstetrician must choose between vaginal birth, at the risk of the life of the mother and the fetus, or a C-section, which is safer but requires a longer recovery process. In Table 21.7, we model the stage at which a fully informed prospective mother would decline a C-section, but just barely; with a slight increase in the risk of harm to her or the fetus, the mother would opt for a C-section.

Table 21.7 Vaginal Birth and C-section									
EU				> _{UTIL} < _{IntraPP}	EU				
2/3	0	1	1		2/3 - ε	2/3 - ε	2/3 - ε	2/3 - ε	
2/3	1	0	1		2/3 - ε	2/3 - ε	2/3 - ε	2/3 - ε	
2/3	1	1	0		2/3 - ε	2/3 - ε	2/3 - ε	2/3 - ε	
Vaginal Birth					C-section				

⁴ A similar case and analysis is presented in Fleurbaey and Voorhoeve (2013).

For example, suppose that the mother has a $2/3$ chance of the best outcome (say a successful vaginal birth) at utility 1 and a $1/3$ chance of the worst outcome (say, death) at utility 0. If the mother chooses a C-section, she is certain to end up with utility $2/3 - \epsilon$.

Given that this concerns a pattern of decisions followed by obstetricians, we imagine that multiple women are in this situation.

Because the expectation of vaginal birth is greater than of C-section, the affected prospective mothers will choose for a vaginal birth. If the prospective mother is herself an MD, then the obstetrician will simply exercise fiduciary care and execute her will.

$$V_{UTIL}(VaginalBirth) > V_{UTIL}(CSection) \quad (14)$$

If the prospective mother is not an MD, then the obstetrician will be deciding for her. Now suppose that the obstetrician is sensitive to the intra-personal prospect distribution. For a sufficiently large value of the parameter β_{CE} and a sufficiently small ϵ ,

$$V_{\beta_{CE}=x>0}(VaginalBirth) < V_{\beta_{CE}=x>0}(CSection) \quad (15)$$

When deciding for non-MD prospective mothers, they are deciding for others. Even though it may be reasonable for the prospective mother herself to take the risk of a vaginal birth, the obstetrician wishes to play it safe. In this case, playing it safe is playing it a bit safer than a reasonable and fully informed prospective mother would choose to play it herself. This can be modeled by assigning a value greater than 0 to the certainty equivalent parameter β_{CE} , measuring the decision-maker's sensitivity to the intra-personal prospect distribution.⁵

21.4.5 Summing Up

The first four dilemmas nicely map onto our four distributional sensitivities. The following Table 21.8 provides an overview.

In each dilemma, it is a particular sensitivity that provides a counter weight to the utilitarian recommendation. Someone who has this sensitivity will reasonably substitute the value of the prospect in the table for the utilitarian value of the prospect.

We now move on to the dilemma of *Saving Few at Greater Risk to Many*, which is slightly more complex because it incorporates multiple distributional sensitivities.

⁵ Lara Buchak (2017: 22–4) argues that in making risky decisions for others when not knowing their risk attitude, we should operate with the most risk averse attitude within the range of what is reasonable, but if we know their risk attitude we should operate with their risk attitude. This could provide an alternative explanation. Because the obstetrician and the MD prospective mother have the same medical background, the obstetrician may be presumed to be more knowledgeable of the MD prospective mother's risk attitude. If the prospective mother is not an MD, then the obstetrician is less knowledgeable and decides for her with a more risk-averse attitude.

Table 21.8 Overview

Dilemma	Sensitivity for Distribution	Value of the Prospect
Restrictive Intervention	Inter-Personal Prospect	Equally Distributed Equivalent of the Personal-Prospect Expectations
Diversification	Inter-State	Certainty Equivalent of the State Averages
Unwelcome Risk Reductions	Intra-State	Expectation of the State Equally Distributed Equivalents
Not on My Watch	Intra-Personal Prospect	Average of the Personal-Prospect Certainty Equivalents

21.5 SAVING FEW AT GREATER RISK TO MANY

21.5.1 Cases

Here are five cases of in which we must choose between saving the few or reducing the risk to the many.

- (i) We can either provide expensive anti-retrovirals to AIDS patients or increase efforts to prevent HIV transmission (for example, through condom dispersal programs).
- (ii) We can either treat those infected with Ebola or we can invest in prevention through vaccination projects.
- (iii) We can either invest millions on saving a miner who is trapped or direct the funds to improving mine safety so as to prevent future accidents.
- (iv) In the biblical Parable of the Lost Sheep (Luke 15: 3–7) we can go out and save one sheep while leaving the 99 sheep in the barn at the risk of a visit by wolves or we can protect the 99 sheep.⁶
- (v) Diesel engines emit more NO_x and particulates. This imposes risk on present people who suffer from respiratory diseases. Fuel engines emit more CO₂ and impose risks on future people due to climate change, and these risks will affect many. We can protect the present few by policies favorable to fuel engines, or we can reduce the risk to the future many by policies favorable to diesel engines.

There is a question as to whether the many are facing more or less correlated risk. The stories can be filled in in different ways. For some diseases that we fail to prevent,

⁶ Luke 15 is a response to the Pharisees who object that Jesus keeps the company of sinners and tax collectors. Jesus responds that there is more joy in saving people who have gone astray than in attending to the righteous. My interpretation may be a bit of a stretch as biblical scholarship goes, but one could say that the cost of saving a few sinners is that one cannot give due attention to the righteous who are thereby put at increased risk of sinning.

the risks to the many are independent. For other diseases, the risks are correlated; that is, there is the risk of an outbreak. Investment in mine safety may address more or less correlated risk. Depending on the hunting behavior of the wolves, the risk to the sheep in the barn may be more or less correlated. In the case of climate change risks are clearly correlated.

All of these cases have the structure of a dilemma. In the case of independent risk, the pull comes from two directions. We want to save the most expected lives, yet it seems heartless not to save the few. In the case of correlated risk, there is an additional pull from a third direction, viz. we are wary of catastrophes in which things go wrong for the many.⁷

21.5.2 Modeling Saving Few at Greater Independent Risk to Many

Suppose that person n is sure to die unless we save them. If we forego saving them, then we can increase the survival chance for persons $1, \dots, n-1$ from 0.95 to $0.95 + \epsilon$. Let us assume that these survival chances are independent. Under *Treatment*, resources are being allocated to person n who will now survive whereas otherwise they were sure to die. Under *Prevention*, we let person n die, but we reduce the risk to persons 1 through $n-1$, who now face a better prospect. Suppose that $(n-1)\epsilon > 1$ so that the utilitarian decision-maker prefers *Prevention* to *Treatment* (see Table 21.9).

However, if a decision-maker is sensitive to the distribution of the expectations (that is, if they are sensitive to the inter-personal prospect distribution), then they will prefer *Treatment* to *Prevention*.

Table 21.9 *Prevention and Treatment*

<table><tr><td>EU</td><td></td><td></td><td></td><td></td></tr><tr><td>0.95 + ε</td><td>0</td><td>1</td><td>1</td><td>...</td></tr><tr><td>...</td><td>1</td><td>0</td><td>1</td><td>...</td></tr><tr><td>...</td><td>...</td><td>...</td><td>...</td><td>...</td></tr><tr><td>0.95 + ε</td><td>1</td><td>1</td><td>0</td><td>...</td></tr><tr><td>0</td><td>0</td><td>0</td><td>0</td><td>...</td></tr></table>	EU					0.95 + ε	0	1	1	...	...	1	0	1	...	...	...	...	...	...	0.95 + ε	1	1	0	...	0	0	0	0	...	> UTIL < InterPP	<table><tr><td>EU</td><td></td><td></td><td></td><td></td></tr><tr><td>0.95</td><td>0</td><td>1</td><td>1</td><td>...</td></tr><tr><td></td><td>1</td><td>0</td><td>1</td><td>...</td></tr><tr><td>...</td><td>...</td><td>...</td><td>...</td><td>...</td></tr><tr><td>0.95</td><td>1</td><td>1</td><td>0</td><td>...</td></tr><tr><td>1</td><td>1</td><td>1</td><td>1</td><td>...</td></tr></table>	EU					0.95	0	1	1	...		1	0	1	...	...	...	...	...	...	0.95	1	1	0	...	1	1	1	1	...
EU																																																														
0.95 + ε	0	1	1	...																																																										
...	1	0	1	...																																																										
...	...	...	...	...																																																										
0.95 + ε	1	1	0	...																																																										
0	0	0	0	...																																																										
EU																																																														
0.95	0	1	1	...																																																										
	1	0	1	...																																																										
...	...	...	...	...																																																										
0.95	1	1	0	...																																																										
1	1	1	1	...																																																										
Prevention		Treatment																																																												

⁷ The problem of saving few at the cost of a greater risk to many is discussed by Schelling (1968), Reibetanz (1998), Brock and Wikler (2009), Otsuka (2015), and Frick (2015a, 2015b). The Lost Sheep case is discussed in Bovens and Fleurbaey (2012) and Dieselgate is discussed in Bovens (2017).

Formally, for the utilitarian decision-maker the value of the prospects is such that:

$$V_{UTIL}(Prevention) > V_{UTIL}(Treatment) \quad (16)$$

However, for sufficiently large α_{EDEE} and sufficiently small n and ϵ , the value of the prospects is such that:

$$V_{\alpha_{EDEE}=x>0}(Treatment) > V_{\alpha_{EDEE}=x>0}(Prevention) \quad (17)$$

Exactly the same reasoning applies here as in the case of *Restrictive Intervention*, discussed in Section 21.4.1.

21.5.3 Modeling Saving Few at Greater Correlated Risk to Many

In Table 21.10, we fill in the story of the Lost Sheep as a case of correlated risk. We call the action of saving the one sheep “*Saving One Sheep*” and attending to the sheep in the barn “*Protecting 99 Sheep*.” *Saving One Sheep* saves the lost sheep on row 100, but at the expense of leaving the sheep on rows 1 through 99 and subjecting them to the risk of a visit from a pack of wolves. *Protecting 99 Sheep* gives up on the sheep on row 100 but makes sure that the sheep on rows 1 through 99 are properly protected.

If the survival chances of the 99 sheep in the barn are $(99 - 1)/99 = 0.989 \dots$, then the utilitarian decision-maker is indifferent:

$$V_{UTIL}(Protecting) = V_{UTIL}(Saving) \quad (18)$$

A decision-maker who is sensitive to the inter-personal prospect distribution objects to the fact that *Protecting 99 Sheep* imposes a wider spread of expectations on the sheep.

$$V_{\alpha_{EDEE}=x>0}(Protecting) < V_{\alpha_{EDEE}=x>0}(Saving) \quad (19)$$

Table 21.10 *Protecting 99 Sheep and Saving One Sheep*

<table><tr><td>EU</td><td></td><td></td><td></td></tr><tr><td>1</td><td>1</td><td>1</td><td>1</td></tr><tr><td>...</td><td>...</td><td>...</td><td>...</td></tr><tr><td>1</td><td>1</td><td>1</td><td>1</td></tr><tr><td>0</td><td>0</td><td>0</td><td>0</td></tr><tr><td>SA</td><td>0.99</td><td>...</td><td>0.99</td></tr></table>				EU				1	1	1	1	...	...	...	...	1	1	1	1	0	0	0	0	SA	0.99	...	0.99	$\succ_{InterState}$ $\prec_{InterPP}$	<table><tr><td>EU</td><td></td><td></td><td></td><td></td></tr><tr><td>0.989...</td><td>1</td><td>...</td><td>1</td><td>0</td></tr><tr><td>...</td><td>...</td><td>...</td><td>...</td><td>...</td></tr><tr><td>0.989...</td><td>1</td><td>...</td><td>1</td><td>0</td></tr><tr><td>1</td><td>1</td><td>...</td><td>1</td><td>1</td></tr><tr><td>SA</td><td>1</td><td>...</td><td>1</td><td>0.01</td></tr></table>					EU					0.989...	1	...	1	0	...	...	...	...	...	0.989...	1	...	1	0	1	1	...	1	1	SA	1	...	1	0.01
EU																																																															
1	1	1	1																																																												
...	...	...	...																																																												
1	1	1	1																																																												
0	0	0	0																																																												
SA	0.99	...	0.99																																																												
EU																																																															
0.989...	1	...	1	0																																																											
...	...	...	...	...																																																											
0.989...	1	...	1	0																																																											
1	1	...	1	1																																																											
SA	1	...	1	0.01																																																											
Protecting 99 Sheep				Saving One Sheep																																																											

A decision-maker who is sensitive to the inter-state distribution objects to the fact that *Saving One Sheep* leaves a wider spread between the average state utilities: Things may just go horribly wrong in *Saving One Sheep*.

$$V_{\beta_{CEA}=x>0}(\textit{Protecting}) > V_{\beta_{CEA}=x>0}(\textit{Saving}) \quad (20)$$

There is an intuitive story to be told here. On the one hand, we may be especially sensitive to the fact that some individuals have very low expectations. In this case we do not want to just give up on the sheep on row 100. Sensitivity to the inter-personal prospect distribution makes us more eager to save the few than a utilitarian decision-maker. On the other hand, we may particularly care about the possibility that things could just go horribly wrong for the collective. In this case, we should be prepared to give up on row 100 in order to avoid the possibility of a grand downfall for all. Sensitivity to the inter-state distribution makes us more eager to reduce the risk to the many than a utilitarian decision-maker.

So, in the case of correlated risk, the relevant sensitivities pull in opposite directions from the utilitarian ranking.

21.6 CRITICS

Each of these sensitivities has had its detractors over the years, and debates in the literature about making risky decisions for others can be captured in terms of how much weight these sensitivities should receive in particular situations.

21.6.1 Against Inter-Personal Prospect Distribution Sensitivity

Table 21.11 is a variant of a case presented by Reibetanz (1998). Suppose that 100 farmers will be ploughing a field. There is a bomb in the field and one is bound to blow themselves

Table 21.11 *Safe Detonation and Lethal Bomb*

<table><tr><td>1</td><td>1</td><td>1</td><td>1</td></tr><tr><td></td><td>1</td><td>1</td><td>1</td></tr><tr><td>...</td><td>...</td><td>...</td><td>...</td></tr><tr><td>1</td><td>1</td><td>1</td><td>1</td></tr><tr><td>0.9</td><td>0.9</td><td>0.9</td><td>0.9</td></tr><tr><td>Average Expectation: 100.9/101</td><td></td><td></td><td></td></tr></table>				1	1	1	1		1	1	1	...	...	...	...	1	1	1	1	0.9	0.9	0.9	0.9	Average Expectation: 100.9/101				>_{UTIL} <_{InterPP}	<table><tr><td>0.99</td><td>0</td><td>...</td><td>1</td><td>1</td></tr><tr><td></td><td></td><td></td><td></td><td></td></tr><tr><td>...</td><td>...</td><td>...</td><td>...</td><td>...</td></tr><tr><td>0.99</td><td>1</td><td>...</td><td>1</td><td>0</td></tr><tr><td>1</td><td>1</td><td>...</td><td>1</td><td>1</td></tr><tr><td>Average Expectation: 100/101</td><td></td><td></td><td></td><td></td></tr></table>					0.99	0	...	1	1						...	...	...	...	...	0.99	1	...	1	0	1	1	...	1	1	Average Expectation: 100/101				
1	1	1	1																																																												
	1	1	1																																																												
...	...	...	...																																																												
1	1	1	1																																																												
0.9	0.9	0.9	0.9																																																												
Average Expectation: 100.9/101																																																															
0.99	0	...	1	1																																																											
...	...	...	...	...																																																											
0.99	1	...	1	0																																																											
1	1	...	1	1																																																											
Average Expectation: 100/101																																																															
Safe Detonation				Lethal Bomb																																																											

up in doing so. A uniquely qualified bomb expert can detonate the bomb safely but in so doing will incur pneumonia. We are comparing the following distributions. Utility 1 stands for full health, 0 stands for death, and 0.9 for incurring pneumonia.

Safe Detonation, which has the greater utilitarian value, indeed seems to be the morally preferable option. However, if we are sufficiently sensitive to the inter-personal prospect distribution, that is for a sufficiently high α_{EDEE} , then the value of *Lethal Bomb* exceeds the value of *Safe Detonation*.

In this particular case, there are indeed various features that would make one hesitant about supporting the solution that is sensitive to the inter-personal prospect distribution:

- (i) The harm to the expert (pneumonia) is substantially smaller than the harm that is bound to befall one of the farmers (death).
- (ii) The risks are anti-correlated and the harm is sure to occur (unlike for independent risk).
- (iii) We are making a recommendation about what the expert should do, not about the policy that a decision-maker should implement affecting others.

Suppose that the expert is bound to lose a limb in carrying out their task, with the utility of this health state set at 0.50 (*Harmful Detonation*), or that they run a 0.50 chance of losing their lives (*Risky Detonation*). Suppose that the risks to the farmers are independent so that there is a small chance that nobody will be affected. Let us call this *Lethal Bomb**. And suppose that we are a decision-maker who needs to instruct the expert what to do. In this case, the utilitarian calculus would still recommend *Harmful Detonation* and *Risky Detonation* over *Lethal Bomb**, because its average expectation equals $100.5/101 > 100/101$. But now, I submit, we would be much less comfortable following this recommendation.

21.6.2 Against Inter-State Distribution Sensitivity

Keeney (1980) studies prospects with binary utilities—say, either the people in the prospect will live or they will die. He shows that there is a conflict between the goal of risk equity, that is, the goal of distributing risk equally between individuals, and the goal of catastrophe avoidance, that is, minimizing the risk that a large number of individuals will be hit. Indeed, the best way to avoid a catastrophe is to concentrate the risk on a few individuals, but this sacrifice of a small group is anti-egalitarian.

Table 21.12 represents a case of independent risk. Compare the following two prospects with binary utilities. Suppose that one person is bound to die and the other is bound to live; call this prospect *One Lives One Dies*. We can spread the risk so that each has an independent chance of dying of 0.50; call this prospect *Coin Flip*. The certainty equivalent of the state averages of *One Lives One Dies* is 0.50. But for $\beta_{CEA} > 0$, the certainty equivalent of the state averages of *Coin Flip* is lower than 0.50. Hence, a decision-maker who is sensitive to the inter-state distribution will try to focus the risk on a single person at the cost of risk equity.

Table 21.12 *One Lives One Dies and Coin Flip*

<table><tr><td>EU</td><td></td><td></td></tr><tr><td>0</td><td>0</td><td>0</td></tr><tr><td>1</td><td>1</td><td>1</td></tr><tr><td>SA</td><td>1/2</td><td>1/2</td></tr></table>	EU			0	0	0	1	1	1	SA	1/2	1/2	$\sim UTIL$ $\succ_{InterState}$	<table><tr><td>EU</td><td></td><td></td><td></td><td></td></tr><tr><td>1/2</td><td>0</td><td>1</td><td>0</td><td>1</td></tr><tr><td>1/2</td><td>0</td><td>0</td><td>1</td><td>1</td></tr><tr><td>SA</td><td>0</td><td>1/2</td><td>1/2</td><td>1</td></tr></table>	EU					1/2	0	1	0	1	1/2	0	0	1	1	SA	0	1/2	1/2	1
EU																																		
0	0	0																																
1	1	1																																
SA	1/2	1/2																																
EU																																		
1/2	0	1	0	1																														
1/2	0	0	1	1																														
SA	0	1/2	1/2	1																														
One Lives One Dies		Coin Flip																																

The same tension holds for correlated risk. Let us revisit the story of the Lost Sheep. The inter-state distribution sensitivity favors *Protecting 99 Sheep* over *Saving One Sheep*. *Protecting 99 Sheep* offloads the risk on one sheep, viz. the lost sheep that will not be saved. Risk equity is embodied in the inter-personal prospect distribution sensitivity which, as we saw, favors *Saving One Sheep*, because it spreads the risk from the lost sheep to the 99 sheep in the barn.

21.6.3 Against Intra-State Distribution Sensitivity

Adler (2012: 523–4) argues that a decision-maker who is sensitive to the intra-state distribution does not satisfy the Axiom of Weak Separability of Persons.

What does the axiom say? Take two prospects A and B with persons $1, \dots, n$. First, construct prospects A^* and B^* by adding a person who is completely unaffected: Whatever states materializes, they receive a fixed utility u^* . Second, construct prospects $A^\#$ and $B^\#$ by adding a person who is also completely unaffected: Whatever state materializes, they receive a fixed utility $u^\#$. Certainly, the axiom says, whether we add one person with a certain prospect of u^* or another person with a certain prospect of $u^\#$ should not make a difference to the ranking of the prospects.

This axiom is violated if we calculate the value of prospects by means of $\alpha_{EDE} > 0$. Consider the prospects in Table 21.13. Set α_{EDE} at $1/2$. Start with the one-person prospects A and B; add one person with certain utility $u^* = 0$; and one person with certain utility $u^\# = 0.49$. Note how $A^* >_{\alpha_{EDE}=1/2} B^*$ but $B^\# >_{\alpha_{EDE}=1/2} A^\#$, in violation of the Axiom of Weak Separability of Persons.

How problematic is this? The axiom *seems* attractive because we confuse being unaffected with being uninvolved in the project. Indeed, it would be problematic if an uninvolved person could just be tagged onto the prospect and the ordering would flip depending on whether we tag on an uninvolved homeless person or an uninvolved millionaire. Uninvolved people should not be included. However, the fact that one is unaffected by a given policy decision does not mean that one is uninvolved. Involvement is determined by one's place in the social world. A choice between remuneration policies in the company where I am employed may not affect me in that it does not make a

Table 21.13 Adler's Objection

	S1	S2		S1	S2
P1	0	1	P1	0.49	0.49
A			B		

			$> \alpha_{EDE} = 1/2$			
	S1	S2			S1	S2
P1	0	1		P1	0.49	0.49
P2	0	0		P2	0	0
$V_{\alpha_{EDE} = 1/2} (A^*) = 1/8 = 0.1250$				$V_{\alpha_{EDE} = 1/2} (B^*) = 0.1225$		

			$< \alpha_{EDE} = 1/2$			
	S1	S2			S1	S2
P1	0	1		P1	0.49	0.49
P2	0.49	0.49		P2	0.49	0.49
$V_{\alpha_{EDE} = 1/2} (A^\#) = 0.4225$				$V_{\alpha_{EDE} = 1/2} (B^\#) = 0.49$		

difference to *my* income. However, I am involved on grounds of being an employee of the company and the choice between prospects should feature me.

Whether an unaffected person involved in a prospect is richer or poorer does make a difference to the intra-state distribution. We posit a decision-maker who is sensitive to the intra-state distribution. In the choice between A and B this sensitivity does not enter in, because these are one-person prospects. The decision-maker prefers A to B, because it offers a higher expectation. Adding a poor person to the prospect does not make B^* more attractive to the decision-maker and they continue to prefer A^* . Adding a person with the same utility as the person in B introduces equality into $B^\#$, which makes it into a more attractive option than $A^\#$ to our decision-maker.⁸

21.6.4 Against the Intra-Personal Prospect Distribution

Otsuka and Voorhoeve (2018), elaborating on an argument proposed earlier in Otsuka and Voorhoeve (2009), construct the following case, modeled on Nagel (1979). There is a 50/50 chance that your child will turn out to be disabled or not. If they are disabled, then they will be slightly better off in town than in the country. If they are healthy, then

⁸ See also Voorhoeve and Fleurbaey (2016: 943–4).

Table 21.14 Utilities in Otsuka and Voorhoeve's Example

Disabled living in the Country	Disabled living in Town	Healthy living in Town	Healthy living in the Country
0	0.10	0.80	1

they will be much better off in the country than in town. Should you move to town or to the country? We set the utilities as follows:

A single child, if idealized to be rational and well-informed, would choose to move to the country because it yields greater expected welfare. The question is whether a decision-maker choosing for the child could overrule this choice. Otsuka and Voorhoeve believe that a decision-maker should opt for the prospect that maximizes the child's expected welfare. If there are two children in the prospect facing fully correlated risks, then the same holds. But if they are facing fully anti-correlated risk, then we should choose for *Town*. See the rankings for Otsuka and Voorhoeve in Table 21.14.

Otsuka and Voorhoeve are sensitive to the intra-state distribution. They favor *Town* in the case of anti-correlated risk because the health inequalities within each state are less pronounced than in *Country*. However, they are not sensitive to the intra-personal prospect distribution. We can only overrule the choice of the persons in the prospects on grounds of intra-state inequalities.

Parfit (2012) disagrees, arguing that the parents should choose *Town* in each case. His argument is prioritarian: It is better to procure the smaller benefit of $0.1 - 0 = 0.1$ by being located in town for a child that turns out to be disabled than the larger benefit of $1 - 0.8 = 0.2$ by being located in the country for a child that turns out to be healthy. He objects to Otsuka and Voorhoeve that whether risks are correlated or anti-correlated cannot possibly make a difference to the moral ranking over the prospects. The fact that my interests matter more when I am disabled is morally significant as such and not because I am poorly off relative to someone else being better off.

Otsuka and Voorhoeve disagree. In the case of anti-correlation, one sibling will turn out healthy and one sibling will turn out disabled. By moving to the country, we know that a situation will ensue in which the disabled sibling can make the following forceful complaint: "Dear Parents, I have a reasonable complaint about your moving us to the country. It was a benefit to my sibling, but a harm to me, and even though the harm to me is smaller than the benefit to my sibling, you should have given priority to whoever would turn out to be worse off." The same cannot be said in the case of correlated risk.

That is true, but something very similar can be said by the siblings if they both end up being disabled. They may complain: "Dear Parents, we have a reasonable complaint about your moving us to the country. It is a harm to us. Granted, the harm is smaller than the benefit that it would have offered us had we been healthy. However, you should have given more weight in your decision-making to the harm in case we would turn out to be disabled than the benefit in case we would turn out to be healthy." Why would this counterfactual comparison with counterpart siblings in the case of correlated risk

Table 21.15 Otsuka and Voorhoeve's Objection to Prioritarianism

<table><tr><td>Disabled</td><td>Healthy</td></tr><tr><td>0.1</td><td>0.8</td></tr></table>	Disabled	Healthy	0.1	0.8	$\prec_{O\&V}$ $\succ_{Parfit}$ $\succ_{IntraPP}$	<table><tr><td>Disabled</td><td>Healthy</td></tr><tr><td>0</td><td>1</td></tr></table>	Disabled	Healthy	0	1
Disabled	Healthy									
0.1	0.8									
Disabled	Healthy									
0	1									
<i>Town</i>		<i>Country</i>								

<table><tr><td>Disabled</td><td>Healthy</td></tr><tr><td>0.1</td><td>0.8</td></tr><tr><td>0.1</td><td>0.8</td></tr></table>	Disabled	Healthy	0.1	0.8	0.1	0.8	$\prec_{O\&V}$ $\succ_{Parfit}$ $\succ_{IntraPP}$	<table><tr><td>Disabled</td><td>Healthy</td></tr><tr><td>0</td><td>1</td></tr><tr><td>0</td><td>1</td></tr></table>	Disabled	Healthy	0	1	0	1
Disabled	Healthy													
0.1	0.8													
0.1	0.8													
Disabled	Healthy													
0	1													
0	1													
<i>Town</i>		<i>Country</i>												

<table><tr><td>Disabled</td><td>Healthy</td></tr><tr><td>0.1</td><td>0.8</td></tr><tr><td>0.8</td><td>0.1</td></tr></table>	Disabled	Healthy	0.1	0.8	0.8	0.1	$\prec_{O\&V}$ $\succ_{Parfit}$ $\succ_{IntraPP}$	<table><tr><td>Disabled</td><td>Healthy</td></tr><tr><td>0</td><td>1</td></tr><tr><td>1</td><td>0</td></tr></table>	Disabled	Healthy	0	1	1	0
Disabled	Healthy													
0.1	0.8													
0.8	0.1													
Disabled	Healthy													
0	1													
1	0													
<i>Town</i>		<i>Country</i>												

provide any less of a forceful complaint than the actual comparison between siblings in the case of anti-correlated risk?

I argued earlier that it is perfectly acceptable for a decision-maker who is sensitive to the intra-personal prospect distribution to add some risk aversion to the risk attitudes of the persons in the prospect. The parents of our siblings may argue that the risk of moving to the country is a risk that they cannot take for others—they cannot have the child end up disabled in the country *on their watch*. Sensitivity to the intra-personal prospect distribution can justify a choice of *Town* over *Country* in each of the three cases of Table 21.15. Hence, in this case, sensitivity to the intra-personal prospect distribution could support the same rankings as Parfit's prioritarian argument.

21.7 DISCUSSION

We have discussed four distributional sensitivities that may, in certain decision situations, upset the utilitarian ordering. Multiple sensitivities can enter in at the same time and either reinforce each other or pull in opposite directions. In Bovens (2015a, 2015b), I proposed a model that integrates these sensitivities.

It is a mistake to think that there is a single model that will tell us once and for all how we should resolve dilemmas of risky decision-making for others. The information that is contained in prospects is too impoverished to determine a unique ranking. To determine

**Table 21.16 Sen's
Libertarian Paradox**

Prude	Lewd
O	P
P	L
L	O

a moral ranking over risky prospects affecting others, we need to have access to much more detail than matrices of utilities with persons in the rows, states in the columns, and probability distributions over states.

The situation is comparable to Sen's (1970, 1976) and Gibbard's (1974) discussion of the Libertarian Paradox. In Sen's version of the paradox, Prude and Lewd find a copy of *Lady Chatterley's Lover* in the book store. Lewd thinks it would be a shame for the book to remain unread. She would like to read it herself. But she prefers that Prude read it, because it would do him some good. Prude would rather have it that nobody read the book. But if someone is going to read it, then it better be him, because heaven knows what it might do to Lewd. Table 21.16 contains the preference orderings.

Who, if anyone, should walk out with the book? Again, we have a dilemma. The Principle of Minimal Libertarianism says that one should be decisive about at least some options in one's private sphere, such as, whether to read or not to read a particular book. Prude and Lewd should each be decisive as to whether they will or will not read the book. Hence, $L > o$ and $o > P$ and so, by Minimal Libertarianism, we should hand the book to Lewd. But note that both Lewd and Prude prefer $P > L$. So, by the Pareto Principle, we should hand the book to Prude. What should we do?

Gibbard presents a version of Gilbert and Sullivan's *Trial by Jury* that has a similar structure. It is not quite respectful of modern sensitivities, but as a comedy and a critique of mores, it still passes muster. Edwin would rather not marry Angelina, but if she is going to marry the Judge, then he would rather marry her than see her marry the Judge. Angelina would rather marry Edwin than the Judge, but she prefers marrying the Judge to not marrying anyone. Table 21.17 contains the preference orderings.

Who, if anyone, should get married to whom? Minimal Libertarianism stipulates the following. If the Judge is a willing party, then it is in Angelina's private sphere whether to

**Table 21.17 Gibbard's version
of the Libertarian Paradox**

Edwin	Angelina
O	E
E	J
J	O

marry him or not. If Angelina is a willing party, then it is in Edwin's private sphere whether to marry her or not. So, $J > o$ and $o > E$. By Minimal Libertarianism, we recommend that the Judge marry Angelina. However, both Edwin and Angelina prefer to be married to each other than for Angelina to be married to the Judge. Hence, by Pareto, $E > J$. What should be our recommendation?

In Sen's version, there is something to be said for discounting nosy preferences and respecting each person's desire to read or not read. At least, this is how Sen sees the case. Hence, Minimal Libertarianism wins and the book goes to Lewd. In Gibbard's version, there is something to be said for letting the parties express their preferences. Angelina can threaten Edwin to marry the Judge, and Edwin will come to realize that he can only avoid this outcome by marrying Angelina. Hence, Pareto wins and Angelina and Edwin will tie the knot. At least, this is how Gibbard sees the case.

In Sen's response to Gibbard, he concedes that there is no single solution to the Libertarian Paradox. To resolve the dilemma in a morally sensitive matter, we need much more information than that contained in the preference profile. The preference profile is too informationally impoverished to determine whether Pareto or Minimal Libertarianism should have the overhand.

Similarly, prospects alone are too informationally impoverished to determine a unique ordering. Different distributional sensitivities are fitting in different contexts. And just as in Sen's and Gibbard's Libertarian Paradoxes, different dilemmas may be represented by the same prospects and yet demand very different solutions.

Here is an example of very different types of decisions that can be modeled by the same pair of prospects. Let us revisit the case of screening for breast cancer that we modeled in Table 21.6. Public health dilemmas over alcohol policy can be modeled by the same pair of prospects. There are various things one can do to reduce alcohol consumption—increased taxes, restrictions on deep discounting (that is, a sale price below VAT and excise taxes), minimal unit pricing, limiting sales times, and so forth. Many people object to restrictive alcohol policies. However, when we consider the number of alcohol-related deaths prevented in the population at large by restrictive policies, one might think that this annoyance is a small price to pay for the benefits gained. The stylized prospects for *No Screening* and *Screening* could also serve as models for respectively a liberal alcohol policy and a restrictive alcohol policy.

A decision-maker with a sensitivity for the intra-state distribution might defend a policy of intensive screening for breast cancer, overruling the preferences of a majority of the women involved. We know that it is highly likely that many more women would die with less intensive screening, and from a public health standpoint saving these women through intensive screening is worth the annoyance and unnecessary treatment. I have some sympathy for this position. However, in the case of alcohol policy, I am less sympathetic to restrictive alcohol policies. Certainly, we can save many lives by introducing more restrictive alcohol policies, and yet, I am less inclined to overrule the preferences of a population that resists such moves. In England, suggestions of more restrictive alcohol policies are met with the defiant response: "We do not want to become Sweden."

So even though the pair of prospects that model both public health dilemmas are the same, I am more inclined to let the Pareto condition be overruled in the case of breast cancer screening than in the case of alcohol policies. What accounts for the difference? In the case of breast cancer screening, who will and will not be affected by breast cancer is down to brute luck. But in the case of alcohol policies, who will and will not be affected by alcohol-related morbidity is to a certain extent down to the choices that people make. For this reason, I am less willing to overrule the preference of many currently responsible drinkers who want to continue enjoying their pint English-style. This distinction is not reflected in prospects: All we have is a probability distribution over better or worse consequences that may come about under different screening regimes or different alcohol policies. Just as in the Sen-Gibbard debate, there is information that is extraneous to the formal model, viz. whether people have more or less control over health outcomes, that determines how much weight conflicting reasons merit in resolving the dilemma.

21.8 CONCLUSION

What is the good of modeling, if not to determine a single decision algorithm ideally backed up by an axiomatic justification? What progress did we make by modeling these multiple sensitivities? Making risky decisions for others in a morally responsible matter is not an algorithmic exercise and the formal machinery is not meant to make it such. However, formal machinery helps us understand what forces are at work when we find ourselves in a quandary and what is driving people when they disagree in the forum. This is an important contribution to moral decision-making. It is a mistake to expect more definite answers from modeling.

ACKNOWLEDGMENTS

I am grateful for comments on earlier drafts by David Kinney, Ruth Tomlin, Alex Voorhoeve, and Mark White.

REFERENCES

- Adler, Matthew. 2012. *Well-Being and Fair Distribution: Beyond Cost-Benefit Analysis*. Oxford: Oxford University Press.
- Aiken, Alexander M., Calum Davey, James R. Hargreaves, and Richard J. Hayes. 2015. "Re-Analysis of Health and Educational Impacts of a School-Based Deworming Programme in Western Kenya: A Pure Replication." *International Journal of Epidemiology* 44: 1572–80.
- Atkinson, Anthony B. 1970. "On the Measurement of Inequality." *Journal of Economic Theory* 2: 244–63.
- Bovens, Luc, and Marc Fleurbaey. 2012. "Evaluating Life or Death Prospects." *Economics and Philosophy* 28: 217–49.

- Bovens, Luc. 2015a. "Concerns for the Poorly Off in Ordering Risky Prospects." *Economics and Philosophy* 31: 397–492.
- Bovens, Luc. 2015b. "Evaluating Risky Prospects: The Distribution View." *Analysis* 75: 243–53.
- Bovens, Luc. 2017. "The Ethics of Dieselgate." *Midwest Studies in Philosophy* 40: 262–83.
- Brock, Dan W., and Daniel Wikler. 2009. "Ethical Challenges in Long-Term Funding For HIV/AIDS." *Health Affairs* 28: 1666–76.
- Buchak, Lara. 2017. "Taking Risks Behind the Veil of Ignorance." *Ethics* 127: 1–35.
- Davey, Calum, Alexander M Aiken, Richard J. Hayes, and James R. Hargreaves. 2015. "Re-Analysis of Health and Educational Impacts of a School-Based Deworming Programme in Western Kenya: A Statistical Replication of a Cluster Quasi-Randomized Stepped-Wedge Trial." *International Journal of Epidemiology* 44: 1581–92.
- Donaldson, David, and John A. Weymark. 1980. "A Single-Parameter Generalization of the Gini Indices of Inequality." *Journal of Economic Theory* 22: 67–86.
- Fleurbaey, Marc. 2010. "Assessing Risky Social Situations." *Journal of Political Economy* 118: 649–80.
- Fleurbaey, Marc, and Alex Voorhoeve. 2013. "Decide as You Would with Full Information! An Argument against ex ante Pareto." In Nir Eyal, Samia Hurst, Ole Norheim and Dan Wikler (eds.), *Inequalities in Health: Concepts, Measurement, and Ethics* (Oxford: Oxford University Press), pp. 113–28.
- Frick, Johann. 2015a. "Contractualism and Social Risk." *Philosophy and Public Affairs* 43: 175–223.
- Frick, Johann. 2015b. "Treatment versus Prevention in the Fight Against HIV/AIDS and the Problem of Identified versus Statistical Lives." In Glenn Cohen, Norman Daniels, and Nir Eyal (eds.), *Identified Versus Statistical Lives: An Interdisciplinary Perspective* (New York: Oxford University Press), pp. 182–202.
- Gauthier, David. 1963. *Practical Reasoning: The Structure and Foundations of Prudential and Moral Arguments and Their Exemplification in Discourse*. Oxford: Clarendon Press.
- Gibbard, Alan. 1974. "A Pareto-Consistent Libertarian Claim." *Journal of Economic Theory* 7: 388–410.
- Greaves, Hilary. 2017. "A Reconsideration of the Harsanyi-Sen-Weymark Debate on Utilitarianism." *Utilitas* 29: 175–213.
- Harsanyi, John. 1955. "Cardinal Welfare, Individualistic Ethics, and Interpersonal Comparisons of Utility." *Journal of Political Economy* 3: 309–21.
- Johnson, Erin M., and M. Marit Rehavi. 2016. "Physicians Treating Physicians: Information and Incentives in Childbirth." *American Economic Journal: Economic Policy* 8: 115–41.
- Keeney, Ralph L. 1980. "Equity and Public Risk." *Operations Research* 28: 527–34.
- Miguel, Edward, and Michael Kremer. 2003. "Worms: Identifying Impacts on Education and Health in the Presence of Treatment Externalities." *Econometrica* 72: 159–217.
- Nagel, Thomas. 1979. "Equality." In *Mortal Questions* (Cambridge: Cambridge University Press), pp. 106–27.
- Otsuka, Michael. 2015. "Risking Life and Limb. How to Discount Harms by Their Improbability." In Glenn Cohen, Norman Daniels, and Nir Eyal (eds.), *Identified Versus Statistical Lives: An Interdisciplinary Perspective* (Oxford: Oxford University Press), pp. 77–93.
- Otsuka, Michael, and Alex Voorhoeve. 2009. "Why It Matters that Some Are Worse Off Than Others: An Argument Against the Priority View." *Philosophy & Public Affairs* 37: 171–99.
- Otsuka, Michael, and Alex Voorhoeve. 2018. "Equality and Priority." In Serena Olsaretti (ed.), *The Oxford Handbook of Distributive Justice* (Oxford: Oxford University Press), pp. 65–85.
- Parfit, Derek. 2012. "Another Defence of the Priority View." *Utilitas* 24: 399–440.

- Rawls, John. 1974–1975. “The Independence of Moral Theory.” In *Collected Papers*, Samuel Freeman (ed.), (Boston, MA: Harvard University Press, 1999), pp. 286–302.
- Reibetanz, Sophia. 1998. “Contractualism and Aggregation.” *Ethics* 108: 296–331.
- Schelling, Thomas. 1968. “The Life You Save May Be Your Own.” In *Choice and Consequence* (Boston, MA: Harvard University Press, 1984), pp. 113–46.
- Sean. 2017. “Deworming Might Have Huge Impact, But Might Have Close to Zero Impact.” *GiveWell Blog*, at <https://blog.givewell.org/2016/07/26/deworming-might-huge-impact-might-close-zero-impact/>.
- Sen, Amartya. 1970. “The Impossibility of a Paretian Liberal.” *Journal of Political Economy* 78: 152–7.
- Sen, Amartya. 1973. *On Economic Inequality*. Oxford: Clarendon Press.
- Sen, Amartya. 1976. “Liberty, Unanimity and Rights.” *Economica* 43: 217–45.
- Spiegelhalter, David. 2015. “A Visualization of the Information in NHS Breast Cancer Screening Leaflet.” *Understanding Uncertainty*, at <https://understandinguncertainty.org/visualization-information-nhs-breast-cancer-screening-leaflet>.
- Taylor-Robinson, David C., Nicola Maayan, Karla Soares-Weiser, Sarah Donegan, and Paul Garner. 2015. “Deworming Drugs for Soil-Transmitted Intestinal Worms in Children: Effects on Nutritional Indicators, Haemoglobin, And School Performance.” *Cochrane Database of Systematic Reviews*. Available at <http://cochranelibrary-wiley.com/doi/10.1002/14651858.CD000371.pub6/abstract>.
- Ubel, Peter A., Michael L. DeKay, Jonathan Baron, and David A. Asch. 1996. “Cost-Effectiveness Analysis in a Setting of Budget Constraints.” *New England Journal of Medicine* 334: 1174–7.
- Voorhoeve, Alex, and Marc Fleurbaey. 2016. “Priority or Equality for Possible People.” *Ethics* 126: 929–54.
- Weymark, John A. 1991. “A Reconsideration of the Harsanyi-Sen Debate on Utilitarianism.” In Jon Elster and John Roemer (eds.), *Interpersonal Comparisons of Well-Being* (Cambridge: Cambridge University Press), pp. 255–320.
- Williams, Bernard. 1973. “A Critique of Utilitarianism.” In J.J.C. Smart and Bernard Williams, *Utilitarianism: For and Against* (Cambridge: Cambridge University Press), pp. 75–100.

CHAPTER 22

THE TRAGEDY OF ECONOMICS: ON THE NATURE OF ECONOMIC HARM AND THE RESPONSIBILITIES OF ECONOMISTS

GEORGE F. DEMARTINO

22.1 INTRODUCTION

ECONOMISTS cause harm as they strive to do good: that is the tragedy of economics. This claim is a positive description and not a normative indictment of economic practice. John Hicks put it this way: “Under private enterprise, any ordinary change in economic policy involves a change in the price-system, and any change in price benefits those on one side of the market, and damages those on the other” (1939: 706). But economists give insufficient attention to the harm their practice induces—and *that* claim is both positive and normative.

Here is one important example of normative negligence. *Better Living Through Economics* (2012) is a compilation of essays by prominent economists, edited by long-time American Economic Association Executive Secretary John Siegfried, which celebrates the contributions that economists have made across policy areas. One might expect that a book that takes stock of economists’ impacts in the world would present a rigorous comparison of the benefits and costs associated with the policy initiatives surveyed in the book. This is particularly true of trade theory, where the defense of trade liberalization depends on Kaldor-Hicks comparisons of gains and losses and also where recent empirical research demonstrates the extent of the disruptions associated with

liberalization. The book chapter on trade is by Anne Krueger, who claims that trade liberalization in South Korea in the 1960s and subsequently across developing economies generated extraordinary gains. The chapter acknowledges neither losses nor losers.

Objections to this oversight are preempted in the editor's introduction to the book. There we learn that the authors have been given license to steer clear of the debit side of the ledger. If a policy is consistent with the Kaldor-Hicks compensation principle, Siegfried writes,

then, at least in theory, the policy change could be constructed to make some people better off while making no one worse off. Absent such considerations as envy and other relative income issues, the policy change might then be declared "a good thing." Unfortunately, however, matters are not quite so simple. Whether, in fact, compensation actually must be paid, or whether it is sufficient that it could be paid, even though in fact it is not paid, remains controversial. *The chapters in this volume, by and large, ignore such distributional considerations. To do otherwise would drag the analysis into a morass it could hardly escape.* (2012: 2, emphasis added)

That rhetorical maneuver does not startle economists, but what if the editor were introducing a book on better living through engineering or medicine? Would we still tolerate the suppression of any consideration of the downside of professional practice? In comparison with other professions, economists give themselves a pass, avoiding ethical complexities attending their work so that they can get on with it, even when those complexities entail real, serious, sustained, and even, sometimes, irreparable disruption to the lives of others.

Although Hicks recognized the ubiquity of economist-induced harm, he stopped short of advancing an adequate account of economic harm. Indeed, Hicks' and Kaldor's contributions to the new welfare economics of the 1930s set the profession off in the wrong direction, allowing the profession to believe it could settle ethical controversy by solving a math problem.

I have proposed elsewhere the term *econogenic harm* to name what is presently unnamed and therefore too easily repressed: economist-induced harm (DeMartino 2016). The term is inspired by medical ethics, whose term *iatrogenic* harm refers to physician-induced harm. Among the lessons from medical practice is the finding that even the most skilled, prudent, and virtuous physicians can and do cause harm, owing to features of medical practice that are shared by economic practice—and indeed, that arguably are more acutely felt in economics. This is important for a number of reasons. For too long the allegation that economists cause harm has been utilized as a weapon within the profession to cast aspersions on the practice of this or that economist or, more often, on this or that school of economics. *Harm is what other economists do.* An allegation of that sort can only induce defensiveness on the part of the defendant, and therefore obstructs open-minded exploration of the dark side of economics. It would be better to follow the lead of medical practitioners who understand that the risk of harm can be imperfectly managed but cannot be eliminated. The approach in medicine encourages careful examination of the causes of iatrogenic harm and of better practices and protocols in

order to eliminate avoidable harm, engage the patient on the inescapable risks of harm, and prepare for unavoidable harm so as to improve the response and to lessen its impact.

Would acknowledging the tragedy of economics let the profession off the hook for the harm it induces? There is no reason why it should. In medicine we at once recognize the inevitability of the risk of harm but we are still able to hold physicians, clinics, and hospitals morally accountable for inadequate harm management. Recognition of econogenic harm can similarly induce a willingness within the profession and beyond to investigate with open minds the features of economic practice that induce harm, the complex nature of harm and economic harm, and the ethical entailments of harm-inducing economic practice. Like medicine, economics can look to prevent avoidable harm, engage those we purport to help over the risks of harm associated with the initiatives we recommend, prepare communities for possible harms, and respond appropriately—with compassion and respect and not just (hypothetical!) transfer payments—to the harms that arise despite our best efforts.

In what follows I advance a harm-focused approach to economic practice. I explore the causes of econogenic harm; the complex nature of economic harm, where the term refers to all harms that economic interventions induce; the inadequacies in the standard welfarist approach to economic harm; and harm-centric assessment of economic systems, or what I call *economy harm profile analysis*.

22.2 THE SOURCES OF ECONOGENIC HARM

There are two principal causes of econogenic harm (detailed in DeMartino 2013, 2015). Economic interventions typically induce disparate effects across society's members, as Hicks emphasized. Econogenic harm arises whenever economist-advocated interventions that promote aggregate economic interests worsen the absolute situation of some members of the economy or substantially exacerbate inequality. Overzealous economic policy entrepreneurs sometimes claim that their preferred economic strategies benefit all of society's members. Fortunately, the profession knows better. Third-party harm may not be prevalent in very simple economic systems where the lack of specialization works to ensure that an economic innovation has consistent effects across society. Here a new technological or policy innovation may benefit all, even if unequally. Econogenic harm may also be trivial or even absent in the context of narrowly focused interventions, where Pareto-improving adjustments may present themselves. But in complex economic systems marked by a high degree of specialization, confounding webs of relationships transmit the effects of interventions throughout society. The larger the scope of the intervention, the greater the risk. In such cases, even well-designed economic interventions cause harm to some economic actors—and the harm may be substantial. Economists teach the unevenness of economic impacts every day in applied microeconomics. In the standard economic analysis of minimum wage regulation, for instance, some in society benefit from market liberalization while others necessarily suffer economic harm.

In many other professions that risk harming some to benefit others, the extent of the problem is more constrained. In clinical medicine, for instance, the caregiver regularly confers benefits without inducing widespread harm. Iatrogenic harm typically afflicts the agent whose welfare the caregiver seeks to improve, not third parties. In this context—which distinguishes clinical medicine from economic practice—interventions can target the individual, taking into account her particular preferences, interests, and other specifics of her medical situation about which the attending physician can know quite a bit. It is certainly not the norm in medicine (or medical research) to harm some without their consent for the purpose of benefiting others. The field of public health comes closer to economics in terms of third-party harm because it too sometimes imposes interventions that affect large groups of diverse agents. A case in point is the forced isolation of those who have been exposed to a pathogen during an epidemic to prevent infection of the rest of the population. Those affected may be subjected to greater risk of contagion and are denied the opportunity to take precautionary measures that they would prefer (such as fleeing the afflicted region). Another type of case involves triage in situations where resources are insufficient to attend to all needing care.¹ These cases notwithstanding, much of what public health practitioners do generates widely distributed benefits without imposing serious harm on third parties.

A second source of econogenic harm is epistemic insufficiency. Economists face the condition of what I call *irreparable ignorance*. They do not and in principle cannot know enough to prevent unintended and unforeseeable damage. Economic practice entails theorizing today about interventions that will be introduced tomorrow and will induce effects that may very well continue into a distant future. Equally confounding, anticipation of tomorrow's intervention affects agents' behavior today, which in turn alters the initial conditions under which the policy's effects will unfold. Moreover, the policy's effects will unfold in combination with innumerable other changes in the economy and society that cannot be specified in advance with any degree of precision. It was with good reason that Keynes argued that the future is simply unknowable.² In addition, there is the additional problem that policy typically imposes the same "treatment" on agents with diverse and (what are to the economist) inscrutable preferences, interests, and values. Even when predictions of a policy's impact on prices (for example) can be predicted, we still cannot know how the price shifts will affect agents' welfare and agency.

¹ The economist might reply that triage necessarily occurs even in non-crisis situations in the form of allocation of scarce resources. A doctor treating *this* patient cannot at the same time be treating *that* patient.

² And so is the past, and for similar reasons. Causal explanations of past phenomena require liberal reliance on counterfactual reasoning—on beliefs about what would have happened had the intervention under review (the policy that was introduced) not been introduced. This alternative universe is in principle unknowable since it was foreclosed as soon as the intervention was undertaken. For an early introduction to the literature on counterfactual reasoning that remains relevant see Goodman (1947). For a recent examination of causation in the social sciences that incorporates counterfactual reasoning see King et al. (1994: ch. 3).

Some ignorance is in principle repairable. We may come to know next year with some degree of accuracy the GDP in China today. But in practice economists must typically make consequential decisions before repairable ignorance is repaired. For the purposes of today's interventions the fact that something may be known with a high degree of certainty down the road may be immaterial. And so although it would be of academic interest to parse the in-principle unknowable from the in-principle knowable-but-unknown, there is no need to do so here. Both kinds of ignorance entail the risk of unforeseeable consequences, including unintended and unanticipated harm, in the moment that economic interventions must be undertaken. Waiting for all the relevant information to be known would preclude most economic practice, to the detriment of society (or so we presume).

Harm resulting from error—or, more broadly, from the imperative facing the professional to act under conditions of epistemic insufficiency in a world that defies dependable prediction or control—is a universal problem in the professions. Its salience in economics represents a difference in degree rather than in kind. If we theorize a continuum from the professions in which ignorance is manageable or insignificant to those where it is both unmanageable and consequential, we might place basic civil engineering near the benign pole and medical practice at a middling point. In contrast, economics resides in the perilous zone where irreparable ignorance generates risk of severe harm.

Taken together, unevenness of economic impacts and the extent of ignorance in economics imply that the economics profession faces an urgent duty to engage the matter of econogenic harm. So does the fact that econogenic harms, although substantial, are often indirect, postponed, and otherwise obscured. Unlike many other professionals, economists typically do not have to confront the people who are harmed by their practice and they are generally unaware of the extent of and rarely held accountable for the harm (Angner 2006; Ravallion 2009). Moreover, the complexities of the linkages from intervention to effects complicate post-factum assessment of a policy's impact and interfere with the learning that might otherwise occur, both of which promote the proliferation of biases that obstruct clear-headed assessment of economists' practice (Angner 2006).

22.3 ECONOMIC HARM: THE STANDARD APPROACH

Kaldor-Hicks efficiency continues today to inform policy analysis, often operationalized through cost-benefit analysis.³ Hicks described the principle this way: "If *A* is made

³ Kaldor-Hicks and cost-benefit analysis prevail today in applied economics despite the effort of advocates of social welfare functions (SWFs) to displace them. See Adler (forthcoming), as well as his chapter in this volume, for an extraordinarily clear explication and compelling defense of the SWF approach.

so much better off by the change that he could compensate *B* for his loss, and still have something left over, then the reorganisation is an unequivocal improvement” (1941: 111). On first approach the intuition is appealing. But the assumption set that enables Kaldor-Hicks is extensive and, as we will see, problematic. Most importantly, it comprises controversial claims about what humans value and about the nature of harms they suffer (see Boadway and Bruce 1984; Adler 1998, 2012; Kaplow and Shavell 2002; Adler and Posner 2006; and DeMartino 2015, which explores some of the issues in this section in greater detail).

Under Kaldor-Hicks, normative assessment depends on the level of each agent’s welfare (defined on an ordinal scale) and tends to include as arguments in each individual’s utility function only her own level of consumption of private and public goods, or the satisfaction of self-regarding preferences (Sen 1987). Harm emerges in this account as a diminution in welfare following a reduction in access to goods that the agent values. In addition, the approach requires that diverse goods be welfare *commensurable* with each other, and that there are no lexicographic orderings, implying that all goods are *substitutable*.⁴ Taken together, these assumptions ensure that all harms are *compensable* through monetary transfers. Without that presumption, of course, Kaldor-Hicks fails as a normative standard. Full compensation for any harm simply requires finding the appropriate magnitude of monetary transfer that will restore the harmed agent to his previous level of welfare. If compensation is made, then the agent is taken to suffer no harm from any loss—indeed, she is theorized as indifferent between her pre- and post-harm situations since compensation implies that she has been made whole. Hence Kaldor-Hicks is sometimes described as a *potential* Pareto test.

Kaldor-Hicks is taken to provide a tractable framework for reaching unequivocal judgments on economic policy decisions that entail harm, one that appears to be only minimally normative. In Hicks’s own view, “this criterion is more useful than any other as a basis on which to establish maxims of sound economic policy” (1941: 111). Like Hicks, many economists believe themselves warranted in advocating policies that induce even substantial harm to some in order to augment aggregate welfare (cf. Kanbur 2003). Moreover, the approach is understood to take adequate account of non-pecuniary as well as pecuniary harms. Kaldor presumed that non-pecuniary harms simply warrant a greater level of compensation than in cases where non-pecuniary harms are absent: in cases involving non-pecuniary harms, “something more than their previous level of money income will be necessary to secure their previous level of enjoyment” (1939: 551).

⁴ Contingent valuation is central to Kaldor-Hicks, just as it is in some approaches to cost-benefit analysis and the Coasean tradition; all are grounded in welfarism. This feature emerges as critical when agents suffer “non-pecuniary” harms, such as increased morbidity and depression and the loss of job satisfaction, status, and social and self-respect that can occur in the context of unemployment (Sen 2000). That said, there are practical, technical, and normative differences between Kaldor-Hicks and cost-benefit analysis that are not relevant to this discussion (see Boadway 1974; Adler and Posner 2006). And unlike Kaldor-Hicks, the Coasean approach entails bargaining that in principle induces actual and not just hypothetical compensation.

22.4 KALDOR-HICKS EFFICIENCY: NORMATIVE DEFICIENCIES

Theoretical economists have raised a host of technical and normative objections to Kaldor-Hicks even while it continues to play a central role in much applied economic policy assessment. The primary normative controversy has to do with potential as opposed to actual compensation. Hicks and Kaldor thought that potential compensation was generally sufficient to permit the endorsement of a policy; most economists follow their lead. In contrast, Amartya Sen ridicules the idea that potential compensation provides adequate justification. “In what sense,” he asks, “is a rise of ‘potential welfare’ of interest to *actual* welfare comparisons? Even if gainers *could* overcompensate the losers, why is that an improvement?” (1979: 24). Other controversies over Kaldor-Hicks entail Kantian concerns (White 2006), disputes over whether welfare is tied to existing as opposed to ideal preferences (Goodin 1986; Adler and Posner 2006) and the related matter of endogenous preferences (Stringham 2001, and below), the substantial normative difference between voluntary Pareto improving transactions and coercive “potential” Pareto improvements (White 2009; Buchanan 1959), and concerns about inequality (Kanbur 2003; Kaplow and Shavell 2002). Many of these have been examined extensively in the literature. Here I raise normative concerns that have not attracted the attention they deserve.

The application of Kaldor-Hicks confronts a daunting philosophical challenge: to ascertain the *moral status* of the harms and benefits a proposed policy will induce. Are harms and benefits—or more precisely, the acts of *harming* and of *not benefitting*—morally equivalent in some unproblematic way? Consider two interventions: one that benefits one agent by x units while harming none, and another that will benefit one agent by $3x$ but harm all others by a total of x . Though the first case induces no harm and the second harms all but one agent, Kaldor-Hicks generates the conclusion that the second policy represents an unequivocal improvement over the first. But this conclusion requires the economist to treat harming and not benefitting as not only morally comparable, but in fact, morally identical (Shiffrin 2012). Doing so transforms an otherwise theoretically daunting comparison of benefits and harms into a comparison of *relative harms*—harm to those who will suffer in the shift from the status quo to a new state of affairs, and the harm to those who, in the event of the perpetuation of the status quo, would have benefitted from the new state of affairs. To put it directly, in this account, not benefitting is theorized as harm.

Several leading legal and moral philosophers who have explored the matter of harm see the matter differently, concluding that not harming must be given moral priority over benefitting. The principle of harm’s priority draws a fairly bright line between the duty not to harm and the duty to benefit others (Shiffrin 2012). In this account not benefitting rarely rises to the level of harming.⁵ The paradigmatic case is the duty to rescue

⁵ Geistfeld’s (2008) exposition of tort law also emphasizes the priority of the plaintiff’s security from rights violations (harm) over the liberty of tortfeasors to pursue their interests (benefit). That said, the priority in torts is relative rather than absolute (see pp. 85–8).

when it can be undertaken with little risk to the rescuer (see Feinberg 1984: ch. 4). Only in cases such as this, where an agent facing no serious risk of harm to herself fails to intervene to prevent a significant deterioration below the baseline condition of another, is not benefitting theorized as harming. Those conditions are not typically met when Kaldor-Hicks is invoked by economists since these cases involve potentially serious harm to some (those who will lose under a Kaldor-Hicks efficient policy adjustment) to advance the condition of other agents beyond their baseline.

Under what circumstances should foregone benefits be treated as equivalent to harm, and when should foregone benefits morally trump the harm to those who will suffer under a proposed policy? What is the economic analogy of the duty to rescue? Answering these questions requires consideration of contextual conditions such as the impact of the harm and benefits on agents' substantive freedoms and the relation of those harms and benefits to their life plans, as well as the relative substantive freedoms of the agents to be harmed and benefitted. Attention must be paid to the question whether those who will be harmed are entitled to be insulated from the harm, and whether those who will benefit are entitled to the benefit. Do agents' claims derive from central rights, or are they grounded in what a fair evaluation of the case would conclude are unsustainable privileges?

A related issue deserves brief mention here. Are there deontological (or other) limits on the Kaldor-Hicks efficient policies that deserve support? If there are such limits, what are they? Medical and public health practitioners and ethicists engage these questions and distinguish between cases where it is and where it is not ethically appropriate to trade off the rights of some for the welfare of others (see DeMartino 2015). Economists invoking Kaldor-Hicks do not, tending instead to displace careful judgment with the formulaic application of a mathematical decision-rule.

22.4.1 Compensable Versus Noncompensable Harm

Equally troubling are the standard assumptions about welfare reducibility and compensability of all harm. We have taken note of the fact that Kaldor-Hicks treats all goods as commensurable and substitutable and, as a consequence, views all harms as reparable and compensable. However, the four concepts are analytically distinct.

In standard welfare economics, all harms fall into the upper-left cell of Table 22.1. Absent commensurability and substitutability, and in the presence of irreparability, the compensation at the heart of Kaldor-Hicks would fail to make the harmed agent whole (lower-right cell). Alternatively, although some harms may prove to be noncompensable in the sense that transfer payments cannot suffice to render the harmed person whole, the harms might be repaired (wholly or in part) through other noncompensatory acts by individuals or society. Acts of *acknowledgment*, such as public or private forms of apology or sympathy, may serve to heal the harmed agent in ways that attempted compensation fails to do (Bouris 2007). Harms suffered in the line of duty (for example, by first responders to crises) or in public service more generally might warrant public expression of gratitude, respect, or honor.

Table 22.1 Compensable vs. noncompensable harm

	Reparable Harm	Irreparable Harm
Commensurable, Substitutable Goods	Compensable Harms	Undefined
Incommensurable, Non-Substitutable Goods	Noncompensable Harms Acknowledgment: Apology, Sympathy, Gratitude, Recognition, Respect, Honor	Noncompensable Harms Remedial measures unavailable

Sometimes, transfer payments occur in the context of public or private acknowledgment; indeed, acknowledgment might even manifest as a monetary transfer. This fact can lead to confusion in the minds of economists who tend not to take sufficient care deciphering the social function of monetary transfers in instances of harm.⁶ Money transfers carry diverse meanings depending on the identities of the giver and receiver, the context in which the transfers occur, and the agents' respective purposes in giving and receiving. Whenever it matters from whom the transfer is given and the social circumstances under which it occurs—whenever knowing the size of the transfer alone provides incomplete information about whether it suffices to do the job for which it is intended—we should be alerted to the fact that the transfer is serving purposes other than compensation.

Apology, sympathy, gratitude, recognition, respect, and honor all reflect an awareness of the fact that not all harms are created equal, and not all are compensable.⁷ Some rights violations, for instance, are generally viewed as too precious to repair with compensation (Rendleman 2002). Alternatively, some harms, such as being dishonored, are not amenable to compensation at all: "Once lost, honor is extraordinarily hard, if not impossible, to regain . . . the very idea of [pricing honor] seems inconsistent with the concept" (McGowan 2010: 589, 591).

Which harms are reparable, and which are not? Which reparable harms are compensable, and which are not? Which compensable harms should be compensated, and which should not? What is to be done when economic interventions potentially threaten or actually do cause noncompensable or irreparable harm, and what authority should the economist have in answering these questions? Questions such as these have been explored at length in tort law, which provides answers that are appropriate to that domain (Geistfeld 2008). Most economists adopting the compensation test have not

⁶ See DeMartino (2015), which argues that Deirdre McCloskey's work is representative of the tendency to view all monetary transfers as "compensation."

⁷ Forensic economists treat the concept of harm much more rigorously. For instance, many take non-pecuniary harms to be noncompensable (Ireland 2010).

taken care to address them systematically. This is particularly true of economists theorizing and advocating policy interventions that induce both harms and benefits.⁸

22.4.2 The Complex Nature of (Economic) Harm

Are all harms best theorized simply as reductions in welfare? Table 22.2 presents a taxonomy of harm, which suffices to make the point that harm is internally complex and manifests in heterogeneous forms that may compound well into the future. Agents can experience physical and mental disabilities as well as psychological or emotional suffering; lose income, wealth, or access to valued goods; be exploited; and suffer impairment in the pursuit of their life plans and a diminution in the capacity for self-respect, creativity, inventiveness, or playfulness (Nussbaum 1992). Harm also manifests as the loss of political efficacy, social respect, meaningful connections with others, and community (Sen 1992; Marglin 2008). Harm comprises both objective and subjective components, and the two may conflict. Harm may arise without causing psychological discomfort or unhappiness, as the literature on “adaptive preferences” makes clear (Elster 1982). The fact that harm can occur behind the backs of those suffering it complicates judgments about the presence, nature, and extent of the harm. It also problematizes the issue of who should be authorized to determine whether harm has occurred and how disabling and indictable is the harm. What role should the agent play in the determination whether she has been harmed, and what role is there in this determination for the economist? Is she to serve as an independent harm expert, or should she largely defer to the judgments of those experiencing the harm?

The concept of harm relates in complex ways to central features of human nature and social existence. Agents suffer harm when their welfare, agency, freedoms, liberties, or valued relationships are undermined. A comprehensive account of harm requires an account of human needs, flourishing, subjectivity (and subjectivity formation), and human connections with others and with their natural and social environment. It also requires an adequate conception of rights (civil, political, economic, social, and cultural), freedoms, and responsibilities.

The impact of a harmful event depends not just on the objective effect of the event, but on agents’ sense of a wide range of factors that contributed to the harm. Does the agent view herself as complicit? Does she view the actors who caused the harm as negligent or otherwise culpable, or does she hold them harmless (Shiffrin 2012)? An athlete suffering a serious injury as a consequence of a hard hit by an opponent but in the

⁸ But see White (2015), which dissents from the standard economic view that the state is generally warranted in imposing external costs on the agents generating them. White argues that external costs resulting from wrongful (rights-violating) behavior should be addressed through tort litigation, while external costs resulting from nonwrongful behavior should not be internalized. An implication for the matter at hand might be that only those economic harms that result from wrongful conduct should be compensated—and that compensation should be addressed under tort litigation. The argument is consistent with White (2006), which rejects welfare-improving Kaldor-Hicks interventions that violate rights.

Table 22.2 A taxonomy of harmed or harmful conditions

Physical:	- Pain - Injury or dismemberment - Loss/diminution of physical or mental capacities - Death - Degradation of the physical environment
Psychological:	- Emotional or psychological suffering; depression - Becoming fearful, insecure, or anxious - Becoming ashamed - Loss of Hope - Erosion of self-respect - Loss of capacity for creativity, playfulness, inventiveness, or fraternal feelings
Economic:	- Loss of income, wealth, or welfare/utility - Loss of access to valued goods - Loss of genuine choice over valued goods - Loss of economic security - Loss of economic opportunities (to do, be, or become) - Loss of economic capacities (e.g., to earn a living) - Loss of control over one's economic activities and practices - Alienation from one's labor, output, or nature - Subjection to exploitation, discrimination, or deprivation
Social:	- Loss of community - Loss of place in community (status, influence, or role as contributor) - Loss of respect, recognition, or honor - Loss of political efficacy - Loss of fraternity or meaningful connections with others - Erosion of social capital
Moral:	Erosion, inversion, and/or collapse of some important ethical or spiritual values, virtues, sensibilities, and norms
Autonomy:	- Adaptive preference formation - Impairment in the pursuit of one's life plans - Treatment as mere means and not also as an end - Destruction of a valued way of life - Constriction of one's capabilities or feasibility set - Exacerbation of personal or systemic threats, risk, or instability - Assault on negative or positive rights/freedoms (coercion) - Denial of opportunity to participate in vitally important social, economic, or cultural processes

context of fair play may feel regret, but the same injury caused by a similar act outside of the arena may induce trauma. In the economic context, does an agent who loses her employment view herself as the victim of an injustice or simply unlucky? Does the agent have available means of responding effectively to her situation? Answers to these questions may depend on whether she alone loses her job or whether an external shock has rendered many in her community unemployed; on the range of services available to her to find new opportunities; and on the depth of social exclusion she encounters as a consequence of being unemployed. And as this discussion suggests, agents buffeted by the same objective harmful event—such as being rendered unemployed—may experience that harm very differently owing to their internal and external resources. Divergent effects across agents complicate the matter of assessing the causal significance of any harm-inducing event, the severity of the consequent harm, the responsibility for the harms suffered, and much else. Can we skirt these complications by assessing the effects that policies are reasonably expected to have on a representative agent rather than on the actual agents who will endure them? Or are we obligated to consider how the trauma is experienced by the real agents populating the economy?

Distinct philosophical and economic frameworks disagree fundamentally about what events cause harm, what effects count as harm, and what harms are normatively indictable. In assessing harm the positive and the normative are inescapably linked (Feinberg 1984). Hicks was therefore wrong in claiming his compensation principle is “a perfectly *objective* test which enables us to discriminate between those reorganisations which improve productive efficiency and those which do not” (1941: 111, emphasis added). Even attending narrowly to economic and not social efficiency, the test requires a raft of normative judgments of the sort that economists are not well trained or willing to make.

22.5 A HARM-CENTRIC APPROACH TO ECONOMIC ASSESSMENT

In promoting the public good economists influence institutions, rules, norms, and practices that bear on what I call an *economy harm profile*.⁹ Economies with distinct institutions, rules, and practices will be characterized by distinct harm profiles. This fact represents a challenge and an opportunity for economists. To the degree that economists can affect the harms that characterize an economy, they have an ethical duty to endeavor to improve the economy harm profile. This requires as a first step investigating the harms that proliferate under existing local, national, and international economic arrangements, and under the alternative arrangements economists advocate, as well as

⁹ This section draws extensively on DeMartino (2018), where I explore economy harm profile analysis in substantially greater detail than I am able to do here.

eliminating gratuitous harms while ameliorating inescapable harms. It also entails exploring which harms they take it to be legitimate (and illegitimate) to impose on some (or all) of society's members or outsiders in pursuit of highly valued goods. I emphasize that economy harm profile analysis provides an additional basis for social assessment that complements but is not intended to replace other normative standards.

An economy harm profile comprises five principal features, each of which encompasses multiple sub-categories (see Box 22.1). They are the 1) *nature*, 2) *productivity*, and 3) *distribution* of prevalent, prevented, and covered harms; 4) *mechanisms* of harm generation and distribution; and 5) *consent* and *coercion* that attend the generation and distribution of harms. The features are internally complex, and each deserves closer attention than I can give it here. For present purposes I limit the discussion to the components of harm profile analysis highlighted in Box 22.1.

Box 22.1 Elements of economy harm profile

1. Nature of Prevalent, Prevented, and Covered Harms
 - a) Prevalence, depth, and risk of
 - reparable* versus *irreparable* harms
 - compensable* versus *noncompensable* harms
 - foreseeable* versus *unforeseeable* harms
 - avoidable* versus *unavoidable* harms
 - b) Nature of harms *prevented*, or *diminished* in frequency/severity
 - c) Nature of *exposed* harms versus nature of harms *insured against* or otherwise *ameliorated*
2. Productivity of Harms
 - a) *Necessary* versus *unnecessary* harms
 - b) *Ethically benign* versus *ethically indictable* harms
3. Distribution of Harms
 - a) *Distribution of winners and losers*
 - in each period, and over the course of successive periods*
 - b) Relative stakes, winning and losing
4. Mechanisms of Harm Generation and Distribution
 - a) Direct harm versus indirect harm
 - b) Fairness of harm-generative arrangements
 - c) *Ability of those causing harm to escape consequences*
 - d) Fragility versus antifragility of economic system
5. Consent versus Coercion
 - a) Extent & intensity of coercion within each particular game comprising the economic system.
 - b) Extent & intensity of coercion to play any particular game
 - c) *Extent & intensity of coercion to play any particular class of games*
 - d) *Extent & intensity of coercion to participate in any particular economic system*

22.5.1 Nature of Prevalent, Prevented, and Covered Harms

Reparable versus irreparable harms, and compensable versus noncompensable harms. It bears repeating that not all harms are created equal. Some harms are well-defined as reparable and compensable, just as standard welfare theory presents them (DeMartino 2015). Contra Kaldor, non-pecuniary damages are often irreparable and noncompensable, as we have seen. This insight bears on economy harm profile analysis. An economy that routinely generates serious irreparable and noncompensable harms to some for the benefit of others must be recognized as distinct from and normatively deficient in this respect in comparison with one whose prevalent harms are reparable and compensable.

Foreseeable versus unforeseeable harms, and avoidable versus unavoidable harms. Economic systems diverge in regard to which harms are foreseeable and which are unforeseeable, as well as which are avoidable and which are unavoidable. Keynes (1936) argued that a highly liquid financial system provides ample credit during normal times but also induces unpredictable financial crises that disrupt commerce and long-term investment. Institutions *inter alia* attempt to convert unforeseeable events into foreseeable events in the hope of increasing their ability to prepare for, take advantage of, and/or ameliorate their effects. Institutions also attempt to convert unavoidable into avoidable events, thereby extending control over the relevant landscape. Up against an uncertain future, however, institutions exert only limited control over harm (Taleb 2007). Efforts to avoid or manage certain kinds of harmful events might fail while increasing the prevalence of other harms. The strategies financial institutions pursue individually to insulate themselves from market volatility might, as Keynes warned, induce systemic risk. It follows that distinct complexes of institutions can be characterized by varying abilities to foresee, control, and avoid distinct forms of economic harm, as well as by the harms they unintentionally generate.

22.5.2 Productivity of Harms

Necessary versus unnecessary harms. Necessary or productive harms are those which are vital to the generation of highly-valued, widely-shared goods that the system conveys. Harms that can be plausibly linked to the realization of valued goals are to be accorded a higher moral status than unnecessary or gratuitous harms, those for which the connection to valued goods is remote or absent.

Economic systems are apt to vary in the degree to which they generate and/or tolerate gratuitous harms. Some economic systems may encompass measures to eliminate or reduce gratuitous harms, while others may demonstrate broad tolerance owing to concerns about the cost of regulation, indifference, and other factors. Economic systems will also differ according to the distinctions drawn between significant harms that warrant regulation and *de minimis* harms that do not. Drawing that distinction entails quantitative, qualitative, and normative judgments. Economic systems that

are alike in important respects, then, may differ substantially in terms of the gratuitous harms that proliferate within them.

Claims about the necessity of harms warrant careful scrutiny. We should resist the presumption of the necessity of harm until we have definitive proof and not just an appealing argument. Moreover, even in the face of a plausible argument and evidence that a harm is necessary, the economist must then inquire whether the good that depends on the harm warrants the harm. In this assessment, reasoned argument must prevail over mechanical decision rules like Kaldor-Hicks that treat goods and harms as commensurable. The simple equation of the good with preference satisfaction and harm with the failure to satisfy preferences will not suffice. “Preferences are the creation of experience and, therefore, also of laws and institutions,” Nussbaum (2001: 192) reminds us (see also Elster 1982). Judgments must be made about which preferences deserve respect and which preference satisfaction failures are cause for concern. Such judgments entail value-driven laundering of preferences before according them weight in any harm calculations (Goodin 1986; Adler and Posner 2006: ch. 5).

Furthermore, the finding that a harm is necessary to achieve a good under prevailing social arrangements cannot be taken as proof that the harm is necessary under alternative available arrangements. This point is fundamental: the *necessity* of harm may be and often is *contextually contingent*, necessary only relative to established economic arrangements. A finding of obligatory harm to achieve a good entails what Nussbaum calls a tragic choice, and should trigger in the minds of the professional what she identifies as the *Hegelian question*: “How can we bring it about that citizens do not face such tragic choices all the time?” (Nussbaum 2001: 187). In a harm-centric approach to social analysis, recognition of the contingent relation between a particular harm and a particular good would prod the economist to search for institutional and policy reforms that render a necessary harm unnecessary. Reforms that eliminate tragic choices—that break the link of dependence between particular goods and harms—serve to generate a more benign economy harm profile.

Ethically benign versus ethically indictable harms. Economic systems vary in the ethical status of the harms they induce. I return to this matter below in the discussion of coercion; here I will note that not all serious harms are ethically indictable and not all legally permitted harms are ethically benign. The first point follows from the nature of any reasonably complex society: we harm each other as a routine feature of living with others (White 2015). We often seek the same goods, positions, rankings, and honors under circumstances where the success of one agent necessarily entails loss to another. Two people competing for the same job cannot both win it; and the harm to the loser might be deep and enduring. Yet, we would not be apt to condemn the victor for his success (provided the contest was fair in salient respects). The harm suffered in this case, even if severe, is ethically benign. Alternatively, economic systems may permit ethically illegitimate harm. The legacy admission system at elite universities in the United States is a case in point (Kahlenberg 2010): it is legal in the United States for private universities to reserve a significant number of places for the children of alumni in the entering class

when making admissions decisions. Nonetheless, the harm to more qualified applicants from less influential families who are excluded from admission is illegitimate on the basis of the most basic and relatively uncontested justice criteria.¹⁰

22.5.3 Distribution of Harms

Distribution of winners and losers in each period, and over the course of successive periods. Economic arrangements vary in terms of the distribution of rewards and penalties. The distribution of the risks of harm in period t_0 can be fairly egalitarian. Alternatively, the risks can be concentrated among a majority to the benefit of a risk-insulated minority, or among a minority to the benefit of an insulated majority. The mechanisms that generate the distribution of harms might entail some combination of luck, merit, privilege, rights, and other factors, and they might encompass some degree or other of randomness. Assessment of the justice of the economic system ought to concern itself *inter alia* with harm distribution in each particular time period.¹¹ The imperative to do so is amplified by the degree to which some of the harms are irreparable or are reparable but not in fact repaired.

The distribution of harms over successive periods deserves particular scrutiny. Economic systems may differ in terms of the serial dependence of harm distributions. In one economic system, being harmed in period t_0 may not predispose the agent to being harmed in periods t_1 , t_2 , and beyond, while another may predispose the loser in one period to lose in successive periods. For instance, systems that generate substantial capabilities inequality are apt to generate not just serial dependence in harm distributions but increasing inequality over time (Sen 2000). This is so in part because capabilities deprivation is apt to lead to income deprivation which in turn amplifies capabilities deprivation. In this case, those harmed in t_0 may expect to suffer equal or greater harms in successive periods. A proposed defense of the serial dependence of harms that emerge in an economic system, then, needs to clear a very high bar before that feature of the system could be adjudged to be ethically legitimate.

22.5.4 Mechanisms of Harm Generation and Distribution

Ability of those causing harm to escape consequences. Economic systems may vary in terms of the degree to which they hold actors accountable for the local and systemic

¹⁰ Legacy admissions in the United States is one of a wide range of practices that together (if not individually) constitute structural violence against the disadvantaged. Structural violence often escapes ethical scrutiny (see Galtung 1969).

¹¹ Predominant normative frameworks would align in this regard, though they differ from each other regarding which mechanisms are just and which are unjust. See Rawls (1971), Nozick (1974), and Sen (1992, 2009) for leading contributions to the debate.

risks they generate. Under one set of arrangements, an actor may have what Taleb (2012, 2018) calls “skin in the game,” where he stands to benefit if all goes well but also suffers harm if things go badly. An aircraft engineer who tests her own experimental designs has skin in the game; so does an entrepreneur who invests his own time, energy and funds in a new venture without prospects of a public bailout. Under alternative arrangements, however, an actor may benefit from the upside of his risky behavior but be insulated from the downside, which is borne by others. Those populating financial institutions deemed too big to fail can take extraordinary risks, profiting from the upside while being insulated from the downside by taxpayer bailouts. Professional economists, too, are typically insulated from the consequences of their errors; they do not lose their jobs when the interventions they advocate yield unintended harm.

Advocates of diverse political and moral philosophies are apt to converge on the matter of harm shifting. There is no refuge in libertarianism, welfarism, or any other normative framework for those lacking skin in the game who gamble with the lives of others. An economic system that permits extensive opportunities for agents to offload risk onto others without their consent while remaining themselves insulated from their imprudence is, in this respect, fundamentally unjust (Taleb 2012; Taleb and Sandis 2015).

22.5.5 Consent versus Coercion

Economic systems vary according to the extent to which they embody consent as opposed to coercion. There is good reason to distinguish between the risks of harm agents consent to and the risks that are imposed upon them.¹² Here I take an approach that examines the various levels at which economic agents may face illicit constraints, from the coercion they face within a particular economic game to the coercion they face to participate in certain economic arenas. It may be the case that voluntarism at one level of agent decision-making co-resides with coercion at other levels.

Extent and intensity of coercion within each particular game comprising the economic system. Complex economic systems comprise many distinct cooperative and competitive arenas (or games) that yield rewards and punishments. Ascertaining the degree to which the overall system entails coercive or consensual harm requires as a first step an investigation of each arena to discern whether *within each arena* participants enjoy extensive voluntarism, or whether, instead, they are coerced in their conduct.

A liberal market is often taken to be the paradigmatic case of voluntarism. Here there is no gun to the head. If it is a liberal market, the terms on which transactors exchange depends on the outcome of bargaining in which agents attend to their own interests. No agent is forced by another to accept unsatisfactory terms. A failure to transact is to be

¹² Coercion is a contested concept, of course. See Wertheimer (1987) for a comprehensive treatment of the daunting questions that must be confronted in any satisfactory account.

taken as further evidence of voluntarism. An agent who cannot secure a fair return on what he offers is free to withdraw from the exchange.

The liberal defense of the market as coercion-free has been rejected by heterodox economists for well over a century. Marx ridiculed the notion of freedom and equality between capital and labor in a capitalist economy. Early institutionalists also emphasized coercion in market economies. John R. Commons argued that workers typically face coercion since they represent the long side of the labor market (see Ramstad 1987). What liberal economists view as an arena free from coercion, then, Marxists and institutionalists identify as a complex of coercive pressures. The goal for Marxists is to eliminate coercion, while Commons and other institutionalists seek to establish laws and regulations that decrease coercive imbalances.

Mapping an economy harm profile requires attention to the presence and degree of coercion within each of the harm-generating games that constitute the economic system. A system that comprises many games that are internally consensual or balance the coercive force of the players, is one that is morally superior to one where the games comprise coercive mechanisms or where the coercive force is deeply unbalanced.

Extent and intensity of coercion to play any particular game. A harm-generating game that affords the participants extensive latitude to make their own decisions—such as what risks to take, what strategies to pursue, and so forth—might yet be indictable if the players are effectively coerced into playing. One way to frame the matter before us is in terms of exit options. An arrangement where agents can desist from participating in a particular economic game that threatens their wellbeing, rights, or other valued goods is to be preferred, morally, over one where they lack the effective freedom to exit it (or the voice necessary to improve it). In the liberal view expressed famously by Friedman (1962), a market economy with a sufficiently large number of firms confronting a sufficiently large number of consumers, suppliers, and laborers achieves voluntarism. In this view any transactions that occur within the market should be judged to be voluntary since buyers and sellers have many others with whom to contract.

Liberal economists tend to presume rather than carefully examine the presence or absence of exit options. In *Capitalism and Freedom* Friedman begins discussion of voluntarism with a stylized account of a simple economy, where “independent households” produce for themselves and exchange output with others. Here, “since the household *always* has the alternative of producing directly for itself, it need not enter into any exchange unless it benefits from it” (1962: 13, emphasis added). A paragraph later he applies the same logic to a modern market economy, neglecting to examine whether the disappearance of the household enterprise, with the means of production and other capacities necessary to engage in production, undermines the condition for effective freedom. He finds, unproblematically, that workers who are unwilling to supply labor at the going market rate can withdraw from the labor market and produce independently. The exit option is presumed rather than demonstrated. In fact, and contrary to Friedman, the absence of exit options in the modern economy may help to explain how it is that so many participants play economic games for which they are so poorly suited and in which they are destined to and do suffer serious, serial harms. To recognize the

harms suffered from a game as consensual, we need to know not only that the game's mechanisms afford all actors substantial "freedom to choose" within the game, but also that the participants freely choose to play it.

Extent and intensity of coercion to play any particular class of games. In assessing consent and coercion we need to investigate whether there are alternative *classes* of games, characterized by internal arrangements that are sufficiently diverse so that agents can avoid those games that threaten the grave hazards that agents seek to escape. It does the agent little good to exit one harm-inducing game only to find that the available options all entail similar harm-generating mechanisms. In this case, the exit option is formal rather than substantive, and consent to play the game is illusory. Participation in a game is genuinely voluntary only if an agent has before her the choice of many distinct classes of games that differ one from the other in salient respects, including the hazard profiles they embody.

Extent and intensity of coercion to participate in any particular economic system. Assessing the extent of freedom to exit a particular class of economic games entails mapping the distinct economic systems that constitute an economy. Until now I have spoken of the two interchangeably, but in fact, as prominent economic geographers are now finding, an economy may comprise multiple economic systems (Gibson-Graham 2006). This has normative importance because an economy that is an internally heterogeneous ecosystem, marked by multiple forms of economic provisioning, would on this account be deemed more consensual and less coercive than one that is an economic monoculture, in which just one form of provisioning dominates or excludes others. It follows, practically, that to the degree that economists influence the economic systems that constitute an economy, they should look to promote a flourishing of economic forms of provisioning rather than use their influence to promote one system that they take to be optimal. The gain in freedom from coercion may more than offset any loss of economic efficiency narrowly defined.

22.6 CONCLUSION

It is in the nature of economic practice that it causes harm. Econogenic harm results from the unevenness of the impact of even well-designed policy interventions and the inherent epistemic constraints presented by irreparable ignorance. It is important to face the fact that economists are in the harm business. *Better Living Through Economics?* Were it only that simple. The strategies that economists celebrate for having promoted economic welfare also induce suffering and sometimes that suffering is irreparable.

The ubiquity of harm presents formidable challenges to the profession, which I have examined here. One is theoretical: to displace the simplistic, reductionist welfare account with a more adequate account of harm that acknowledges its heterogeneity and complexity. An adequate account of harm recognizes that not all harms are reparable or compensable—and that, as a consequence, compensation tests cannot suffice as a harm management decision-rule.

Recognition of the prevalence of harm also presents opportunities for meaningful contributions: helping civil society to identify and eliminate gratuitous harms, prepare for unanticipated and unavoidable harms, ensure an equitable distribution of those harms, hold agents generating risk of harm accountable (in part by ensuring that they have skin in the game), and ensure that agents encounter a broad range of non-coercive economic arenas in which to participate. In short, the profession has the opportunity to theorize and work toward the design of more benign economy harm profiles.

ACKNOWLEDGMENTS

Thanks to Katie Aldrich, Abbey Brown, Lauren Craig, Eugenia Goh, Amanda Hayden, Minsun Ji, Paul Kemp, Patrick Sutherland, Evan Williams, and Michelle Woodruff for wonderful research assistance; and to Mark White for insightful comments on a previous draft.

REFERENCES

- Adler, Matthew D. 1998. "Incommensurability and Cost-Benefit Analysis." *University of Pennsylvania Law Review* 146: 1371–418.
- Adler, Matthew D. 2012. *Well-Being and Fair Distribution: Beyond Cost-Benefit Analysis*. Oxford: Oxford University Press.
- Adler, Matthew D. Forthcoming. *Measuring Social Welfare: An Introduction*. New York: Oxford University Press.
- Adler, Matthew D., and Eric A. Posner. 1999. "Rethinking Cost-Benefit Analysis." *Yale Law Journal* 109: 165–247.
- Adler, Matthew D., and Eric A. Posner. 2006. *New Foundations of Cost-Benefit Analysis*. Cambridge, MA: Harvard University Press.
- Angner, Erik. 2006. "Economists as Experts: Overconfidence in theory and practice, *Journal of Economic Methodology* 13: 1–24.
- Boadway, Robin W. 1974. "The Welfare Foundations of Cost-Benefit Analysis." *Economic Journal* 84: 926–39.
- Boadway, Robin W., and Neil Bruce. 1984. *Welfare Economics*. Oxford: Basil Blackwell.
- Bouris, Erica. 2007. *Complex Political Victims*. Bloomfield, CT: Kumarian Press.
- Buchanan, James M. 1959. "Positive Economics, Welfare Economics, and Political Economy." *Journal of Law and Economics* 2: 124–38.
- DeMartino, George F. 2013. "Epistemic Aspects of Economic Practice and the Need for Professional Economic Ethics." *Review of Social Economy* 71: 166–86.
- DeMartino, George F. 2015. "Harming Irreparably: On Neoliberalism, Kaldor-Hicks, and the Paretian Guarantee." *Review of Social Economy* 73 (4): 315–40.
- DeMartino, George F. 2016. "'Econogenic Harm': On the Nature of and Responsibility for the Harm Economists Do as They Try to Do Good." In George F. DeMartino and Deirdre N. McCloskey (eds.), *The Oxford Handbook of Professional Economic Ethics* (Oxford: Oxford University Press), pp. 71–97.
- DeMartino, George F. 2018. "Econogenic Harm and the Case for 'Economy Harm Profile' Analysis." *New Political Economy*, doi: 10.1080/13563467.2018.1526265.

- Elster, Jon. 1982. "Sour Grapes—Utilitarianism and the Genesis of Wants." In Amartya Sen and Bernard Williams (eds.), *Utilitarianism and Beyond* (Cambridge: Cambridge University Press), pp. 219–38.
- Feinberg, Joel. 1984. *Harm to Others*. Oxford: Oxford University Press.
- Friedman, Milton. 1962. *Capitalism and Freedom*. Chicago, IL: University of Chicago Press.
- Galtung, Johan. 1969. "Violence, Peace, and Peace Research." *Journal of Peace Research* 6: 167–91.
- Geistfeld, Mark A. 2008. *Tort Law: The Essentials*. Austin, TX: Wolters Kluwer.
- Gibson-Graham, J.K. 2006. *A Postcapitalist Politics*. Minneapolis, MN: University of Minnesota Press.
- Goodin, Robert E. 1986. "Laundering Preferences." In Jon Elster and Aanund Hylland (eds.), *Foundations of Social Choice Theory* (Cambridge: Cambridge University Press), pp. 75–102.
- Goodman, Nelson. 1947. "The Problem of Counterfactual Conditionals." *Journal of Philosophy* 44(5): 113–28.
- Hicks, John R. 1939. "The Foundations of Welfare Economics." *Economic Journal* 49: 696–712.
- Hicks, John R. 1941. "The Rehabilitation of Consumers' Surplus." *Review of Economic Studies* 8: 108–16.
- Ireland, Thomas R. 2010. "Different Methods Used to Derive Hedonic Damages in Litigation." *Forensic Rehabilitation and Economics* 3: 67–78.
- Kahlenberg, Richard D. (ed.). 2010. *Affirmative Action for the Rich: Legacy Preferences in College Admissions*. New York: Century Foundation.
- Kaldor, Nicholas. 1939. "Welfare Propositions of Economics and Interpersonal Comparisons of Utility." *Economic Journal* 49: 549–52.
- Kanbur, Ravi. 2003. "Development Economics and the Compensation Principle." *International Social Science Journal* 55: 27–35.
- Kaplow, Louis, and Steven Shavell. 2002. *Fairness Versus Welfare*. Cambridge, MA: Harvard University Press.
- Keynes, John M. 1936. *The General Theory of Employment, Interest and Money*. London: Macmillan.
- King, Gary, Robert O. Keohane, and Sidney Verba. 1994. *Designing Social Inquiry: Scientific Inference in Qualitative Research*. Princeton, NJ: Princeton University Press.
- Marglin, Stephen A. 2008. *The Dismal Science: How Thinking Like an Economist Undermines Community*. Cambridge, MA: Harvard University Press.
- McGowan, David. 2010. "Irreparable Harm." *Lewis & Clark Law Review* 14: 577–96.
- Nozick, Robert. 1974. *Anarchy, State, and Utopia*. Oxford: Blackwell.
- Nussbaum, Martha C. 1992. "Human Functionings and Social Justice: In Defense of Aristotelian Essentialism." *Political Theory* 20: 202–46.
- Nussbaum, Martha C. 2001. "The Costs of Tragedy: Some Moral Limits to Cost-Benefit Analysis." In Matthew D. Adler and Eric A. Posner (eds.), *Cost-Benefit Analysis: Legal, Economic, and Philosophical Perspectives* (Chicago, IL: University of Chicago Press), pp. 169–200.
- Ramstad, Yngve. 1987. "Free Trade Versus Fair Trade: Import Barriers as a Problem of Reasonable Value." *Journal of Economic Issues* 21: 5–32.
- Ravallion, Martin. 2009. "Evaluation in the Practice of Development." *World Bank Observer* 24: 29–53.
- Rawls, John. 1971. *A Theory of Justice*. Cambridge, MA: Harvard University Press.

- Rendleman, Doug. 2002. "Irreparability Resurrected? Does a Recalibrated Irreparable Injury Rule Threaten the Warren Court's Establishment Clause Legacy?" *Washington and Lee Law Review* 59: 1343–406.
- Sen, Amartya. 1979. "The Welfare Basis of Real Income Comparisons: A Survey." *Journal of Economic Literature* 17: 1–45.
- Sen, Amartya. 1987. *On Ethics and Economics*. London: Blackwell.
- Sen, Amartya. 1992. *Inequality Reexamined*. Cambridge, MA: Harvard University Press.
- Sen, Amartya. 2000. *Development as Freedom*. New York: Anchor Books.
- Sen, Amartya. 2009. *The Idea of Justice*. Cambridge, MA: Harvard University Press.
- Shiffrin, Seana V. 2012. "Harm and its Moral Significance." *Legal Theory* 18: 357–98.
- Siegfried, John J. 2012. *Better Living Through Economics*. Cambridge, MA: Harvard University Press.
- Stringham, Edward. 2001. "Kaldor-Hicks Efficiency and the Problem of Central Planning." *Quarterly Journal of Austrian Economics* 4 (2): 41–50.
- Taleb, Nassim N. 2007. *The Black Swan*. New York: Random House.
- Taleb, Nassim N. 2012. *Antifragile: Things That Gain from Disorder*. New York: Random House.
- Taleb, Nassim N. 2018. *Skin in the Game: Hidden Asymmetries in Daily Life*. New York: Penguin Random House.
- Taleb, Nassim N., and Constantine Sandis. 2015. "Skin-in-the-Game Heuristic for Protection Against Tail Events." In George F. DeMartino and Deirdre N. McCloskey (eds.), *The Oxford Handbook of Professional Economic Ethics* (Oxford: Oxford University Press), pp. 13–28.
- Wertheimer, Alan. 1987. *Coercion*. Princeton, NJ: Princeton University Press.
- White, Mark D. 2006. "A Kantian Critique of Neoclassical Law and Economics." *Review of Political Economy* 18: 235–52.
- White, Mark D. 2009. "Pareto, Consent, and Respect for Dignity: A Kantian Perspective." *Review of Social Economy* 67: 49–70.
- White, Mark D. 2015. "On the Relevance of Wrongfulness to the Concept of Externalities." *Æconomia* 5: 313–29.

C. ETHICS IN APPLIED ECONOMICS

CHAPTER 23

ECONOMICS, ETHICS, AND HEALTH INSURANCE

DANIEL M. HAUSMAN

ALTHOUGH some people may be lucky enough never to need health care, even they benefit from knowing that they can access health care if they were to need it, and their expectations for a healthy life are improved by access to health care and by public health measures such as clean water and vaccination. Even if other factors, such as nutrition, housing, education, or social inclusion have a greater overall influence on health than does health care, nothing else matches the drama of medical care rescuing individuals at the precipice of death or permanent injury, or, if their fate cannot be improved, comforting them as they pass over that precipice. Health care is of great social importance, and its funding, structure, and coverage are major policy issues in every country. This is particularly obvious at the moment in the United States, but every health care system is under threat as costs rise and new choices of whom and how to treat present themselves.

The provision of health care has always been an important aspect of societies because it epitomizes the compassionate concerns that individuals have for one another. The way that a society treats those who are most vulnerable says a great deal about that society. Every decent society helps those that are frightened, suffering, disabled and at risk of permanent injury and death and comforts them when no help can be provided. Until recently, medicine could accomplish little, and often all that could be done for ill health was to provide comfort and concern. The circumstances today in most of the world are different. Health care often saves lives or dramatically improves its quality. But this increased capacity is costly, consuming between a tenth and a fifth of the entire economic output of most wealthy societies. How health care is provided matters not only for its contribution to health, but for its effects on solidarity, on prosperity and economic growth, and on individual financial security, dignity, and autonomy.

To assess health care institutions and policies requires ethical reflection. Because health care can be so important to individuals, it is crucial that it be provided in an

ethically acceptable way. Getting the distribution of laundry detergent wrong may result in dirty clothes. Getting the distribution of health care wrong may result in humiliation, suffering, and death.

Ethical reflection is not enough; knowing where we want to arrive does not tell us what road to take. Empirical theories concerning the consequences of alternative policies are also required. The findings of all the social sciences are relevant, but I shall focus on economic analysis. Economic appraisal is crucial, if for no other reason than the mere scale of health care. As such a large part of the economy, health care needs to be efficient. Otherwise, massive resources will be wasted that could have been used to benefit people in many ways. Every worthwhile analysis of health care must be informed by both ethics and economics.

Section 23.1 of this chapter discusses efficiency with respect to health care provision. Efficiency is a crucial value that should govern health care systems. It is not a simple notion. Section 23.2 explains why an unregulated health care market is bound to be inefficient. Section 23.3 argues that efficiency cannot be the sole criterion upon which to assess health care systems, arguing that fairness, compassion, and solidarity are also important. Section 23.4 turns to a fifth important value: freedom or choice. Section 23.5 sketches some of the tradeoffs that must be made in implementing a health care system that is economically feasible and consistent with our values. Section 23.6 concludes.¹

23.1 EFFICIENCY IN HEALTH CARE

In its most general understanding, efficiency is a matter of generating some output with a minimum of inputs, or, equivalently, creating as much output as possible with given inputs. In the case of health care, the inputs consist of the services of health care providers (who may be family, friends, neighbors or even strangers, in addition to members of health care organizations) and a variety of facilities, implements, and drugs. Since inputs are purchased with money (which, of course, represents other goods that the society foregoes in order to purchase health care inputs), it is possible to increase efficiency by lowering the cost of inputs without reducing the quantities of inputs themselves. I shall assume that the health budget is given, and I shall not address questions concerning how much should be spent on health care rather than on other valuable goods and services.

The most important output of the health care system is, of course, health. What constitutes health is contested and complicated, and it is difficult to measure overall or “generic health.” For the purposes of this chapter, I shall rely on the reader’s pre-theoretic understanding and avoid discussing the concept of health and its measurement.² Health

¹ This chapter draws upon chapters 15 and 16 of my 2015 book, *Valuing Health: Freedom, Well-being, and Suffering*, and in several ways supersedes that discussion. See Menzel (2012) for an argument for universal health insurance that complements the analysis in this paper.

² For discussions of the concept of health see particularly Boorse (1977, 1997, 2011); Engelhardt (1974); Wakefield (1992, 1999), Reznick (1987), Nordenfelt (2000), and Hausman (2012a; 2015: chs 2 and 3).

is not the only output of a health care system. Health care also bears on the financial security, independence, and general well-being of individuals as well as on productivity and economic growth. The health care system that is most efficient at promoting health is not necessarily most efficient, once one considers the whole range of outputs of a health care system. If one assumes that making people better off is a good thing, then it is obvious that efficiency is of great importance in any domain that uses many resources. The wasted resources could be producing better health or other things that contribute to good lives.

John Broome (2002) has argued that a health care system should use its resources to promote the well-being of the individuals it serves rather than, more narrowly, to enhance their health. He gives two main arguments for this conclusion. First, the value of health is not separable from well-being. In different circumstances the same state or change of someone's body or mind may have a very different value. Second, he argues that policies in domains of state action must not ignore their cross-domain consequences. For example, education has important effects on health, and the health care system has important effects on economic growth and financial security.

Although how good or bad a health state is for an individual depends on the cultural, physical, and technological environment as well as on that individual's objectives and interests, it is possible to assign a value to a health state that averages over different circumstances or to assign a value to health states with respect to some specified "standard" circumstances. Thus, the inseparability of the value of health does not imply that the health care system cannot take as its objective the promotion of population health. Broome's second argument is more potent. He is clearly right that the health care system should not concern itself only with health. For example, it should take into account the impact of health care on financial security and on the burdens of care that families experience. If its policy is to rest on some sort of algorithm, it needs a common measure of the values of the various outcomes—health and non-health—to which its policies lead. Well-being might appear to be the obvious choice for that common measure.

Despite the cogency of this argument, there are serious problems attached to the proposal to take well-being to be the measure of the relevant health and non-health outcomes of health care policies. First, well-being is a contested notion, which is hard to measure (Griffin 1986; Parfit 1986: Appendix I; Sumner 1996; Scanlon 1998: ch. 3). The most common way is to infer what makes people better off from their preferences, either as expressed in their choices or in direct preference elicitation. But individuals may not be good judges of what makes them better off, and their preferences may be distorted by false beliefs, lack of information, and flaws in deliberation like those documented by behavioral economists. In addition, if individuals are not self-interested, they may prefer what they judge to be worse for themselves. For these reasons, the inference from choice to preference and from preference to well-being is precarious (Hausman 2012b). In addition, to quantify population outcomes, which are aggregates of individual outcomes, one must define a cardinal and interpersonally comparable measure of individual well-being from data concerning choices and preference satisfaction.

The only other way of measuring well-being that has any prominence in the literature focuses on subjective experience, either as constituting or indicating well-being.³ Measurements of current mood appear to have good reliability, but it is questionable whether what they capture can be identified with well-being. Retrospective appraisals, on the other hand, are sensitive to the details of their elicitation and they are sometimes inconsistent with what one would infer from the sequence of an individual's moods. Subjective experience is especially dubious as a measure of the value of health, because some health states diminish consciousness and others, such as depression, systematically alter people's perception of how well their lives are going.

In addition to these practical difficulties in assigning to the health care system the objective of using its resources to maximize population well-being, there is a serious conceptual problem in measuring health by well-being. If one measures how badly diminished someone's health is by their well-being, then it follows that someone has a disability if and only if their health state diminishes their well-being. It follows on this view that if deafness is a significant disability, as it appears to be, then those who are deaf must on average have significantly worse lives than those who can hear. But, as members of the Deaf Community reasonably insist, deafness does not imply a lower level of well-being. In the case of deafness and many other disabilities, individuals are able to construct lives that are every bit as meaningful and satisfying as the lives experienced by individuals who have no disabilities (Lane 2002). This recognition, as important as it is, should not obscure the fact that many careers and important human activities are not accessible to those who have impairments such as an inability to hear or to walk. To avoid the false choice between an indefensible denigration of the lives of the disabled and the equally indefensible denial that an inability to walk or see or hear counts as a diminished health state, one must distinguish the measure of a health state from the level of well-being associated with it. (Moreover, as I will discuss later, if those with disabilities have a lower level of well-being, then there is less value in saving their lives than in saving the lives of those who are fully healthy. If, on the other hand, those who adapt to disabilities are just as well off as those without disabilities, then, apart from the difficulties of adaptation, using resources to prevent or cure disabilities does not increase well-being and is completely cost ineffective. Neither of these implications is a happy one.)

What then is the alternative? I think that the most sensible course is to take the central objective of the health care system to be the promotion of health, while recognizing that some adjustments in policy may be necessary, owing to other important consequences of the health care system. Addressing some of those other consequences, such as providing financial security, appears to be compatible with maximizing health. Others, such as lessening the burdens of personal care or improving educational outcomes may require ad hoc adjustments to the goal of health maximization. What principles should guide these adjustments is a difficult question to which I shall return in the next section and in Section 23.5.

³ See Dolan and Kahneman (2008); Kahneman (1999, 2000a, 2000b), Kahneman and Sugden (2005), Kahneman and Krueger (2006), and Kahneman and Thaler (2006).

In practical terms the inputs that the health care system employs are measured by their monetary cost, and the output is indicated by some summary measure of population health improvement, whether it be a crude measure such as life expectancy or infant mortality, or some more intricate measure such as “quality-adjusted life years” or “disability adjusted life expectancy.”⁴ These measures of both inputs and especially outputs are a precarious basis for assessing the efficiency of a health care system. Even if one had an accurate measure of the value of health, which is unlikely, health depends heavily on other factors besides health care. A health care system may appear to be efficient or inefficient because of factors that have nothing to do with health care. For example, in 2014, life expectancy in France was a year longer than in Germany, despite the fact that per capita health care expenditures in France were nearly 400 euros or about 10 percent less than in Germany.⁵ Should one conclude that France’s health care system is more efficient than Germany’s, or is the difference due to diet, education, climate, or life style? Despite these qualms, huge differences in apparent efficiency surely tell us something about the efficiency of health care systems, especially in countries in which lifestyle is similar, as is the case, for example, in the United States and Canada.

It is safer to attribute the changes in health subsequent to a change in health care to the change in health care than to take the level of health in a society to be a good measure of the efficacy of its health care policies. With respect to health care policy, one can define a reasonably definite measure of efficiency, namely cost effectiveness, with effectiveness understood in terms of some measure of health improvement. If a society employs those health care interventions with the lowest ratio of cost to effectiveness, then it will have employed the most effective health care interventions that its budget makes possible. Measures of effectiveness rely on preference surveys, in which there is wide variance and tremendous sensitivity to how the survey questions are posed. Because these measures are imprecise, cost-effectiveness does not permit fine-grained discriminations among policies.

23.2 IS IT COST-EFFECTIVE TO RELY ON THE MARKET?

The senior senator from my state, Ron Johnson, published an op-ed in the *New York Times* (June 26, 2017) extolling the virtue of markets and advocating relying on them to allocate health care, in which he wrote:

The primary goals of any health care reform should be to restrain (if not lower) costs while improving quality, access and innovation. This is exactly what consumer-driven,

⁴ See for example Dolan (1997), Murray (1996), Murray et al. (2002), Hausman (2015), Murray et al. (2012), and <http://www.healthutilities.com/hui3.htm>.

⁵ See http://ec.europa.eu/eurostat/statistics-explained/index.php/Healthcare_expenditure_statistics and <https://www.cia.gov/library/publications/the-world-factbook/rankorder/2102rank.html>.

free-market competition does in other areas of our economy... Once again, a simple solution is obvious. Loosen up regulations and mandates, so that Americans can choose to purchase insurance that suits their needs and that they can afford.

Perhaps in the future, politicians will not make such jejune references to the virtues of free-market competition, and this chapter will be out of date. I certainly hope so, but I am not optimistic. Knowing just enough economics to understand that markets are wonderfully flexible and efficient ways of allocating many goods and services, politicians will be tempted to suppose that markets can address all problems of allocation. Even without understanding the limitations on markets, one ought to be suspicious of Johnson's "simple solution." If relying on the market were so easy, then it would be bizarre that no affluent country leaves health care to the market, even though these countries rely heavily on the market for the provision of other goods and services. Something is obviously wrong with Johnson's simple solution. Is there some reason why free markets in health care do not work well?

Health care is unlike most other services. It is sometimes urgently needed and extremely expensive. Needs for health care are intermittent and often unpredictable. Unlike other essential needs, such as needs for food and housing, health care needs cannot be met by providing people with cash or vouchers with which to purchase health care and then letting them make their own choices. Much of the time people would have little need for the vouchers, but when expensive emergencies strike, the vouchers would not come close to covering the costs. The only feasible way to provide people with access to expensive health care is via insurance, whether it be public or private. In all affluent societies, insurance pays the substantial health bills. Saving to be able to cover these costs is scarcely feasible and would be extremely inefficient.

Despite its virtues in enabling people to smooth their health care expenditures and to distribute financial risks, insurance brings with it serious problems. First, all insurance is subject to moral hazard (Finkelstein et al. 2015). In the case of *ex ante* moral hazard, individuals take less care to avoid needing medical care in the first place because they will not pay for it. The significance of *ex ante* moral hazard is controversial. How many people will choose to smoke because they know that, if they get lung cancer, they will be able to afford treatment that may prolong their lives for several months? More obviously important is *ex post* moral hazard: once they have a medical need, individuals will be less economical in satisfying it if the expense is covered by health insurance. If it costs nothing to go to the doctor, people go more often. But the discomfort, embarrassment, time, and anxiety attached to a doctor's visit will keep the number of visits from exploding. The effects of moral hazard on the choice of treatments are much greater. Given a choice between an inexpensive treatment and an equally effective but more expensive, safer, and less burdensome treatment, why choose the less expensive treatment if the insurance company is paying the bills?

A half-century ago, Kenneth Arrow (1963) proved (given idealizing assumptions) that the optimal response to the problem of moral hazard requires that individuals pay 100 percent of their medical costs up to some maximum and have complete insurance

after their out-of-pocket expenditures reach that maximum. However, most health insurance policies in the United States combine substantial deductibles with co-pays, as another method to limit the effects of moral hazard. By far the most effective way to cope with moral hazard is via coverage limits built into health insurance policies. If one's insurance company refuses to pay for the more expensive but safer and less burdensome treatment, there is no longer a moral hazard problem. Most of the world's health systems control costs via coverage limits and to a considerable extent eschew deductibles and co-pays. Doing so also greatly reduces administrative overhead. This overhead is considerable. Deductibles and co-pays require administrative staff to process, send, and then collect on bills for the deductibles and co-pays, and they impose on patients the burden of understanding, questioning, and paying those bills. Low deductibles and co-pays may create more costs than they save, while high deductibles and co-pays may undermine the value of insurance for all but dire health problems.

Coverage limits are more problematic in government sponsored insurance than in private insurance, because in the latter people blame their insurance company for heartlessly refusing to pay for an expensive treatment that might help grandma, while in the former it is the government that says, "No." This fact gives a reason to prefer a private health insurance system, because the problems of moral hazard are otherwise similar whether insurance is private or public. But we are far from Senator Johnson's vision of diminishing regulations so that individuals "can choose to purchase insurance that suits their needs." Insurance policies will have complex provisions specifying what will and will not be covered. Given the expenses of underwriting, consumers will have relatively few policies to choose among and are unlikely to find a policy that exactly "suits their needs"—if indeed individuals understand what their insurance needs are.

Another problem with insurance, adverse selection, makes it impossible for individuals to purchase insurance that exactly "suits their needs." Insurance that pays for some specific medical need will, of course, be a better bargain for those who expect to have that need. For example, women who are hoping to become pregnant will be eager to buy affordable insurance that will cover the expenses of pregnancy and child birth. Such insurance will "suit their needs." But if most of the purchasers of pregnancy insurance get pregnant, the premiums for privately purchased pregnancy and childbirth insurance policies will have to be extremely high in order to cover the costs and the insurance company's administrative overhead and profits. With costs that high, there is little point to the insurance company offering or individuals purchasing the insurance, and the market will collapse. If people know what they need and it is cheap, there is no point to purchasing insurance; they are better off self-insuring. If, on the other hand, people know that they will need some specific expensive treatment, then they are not going to be able to purchase insurance that will pay for it, because insurance companies that are not headed for bankruptcy will have taken steps to avoid this adverse selection of just the customers for insurance that the insurance company does *not* want.

Although individuals usually do not know exactly what their needs for health care may be, they often have some evidence, and the pool of those purchasing health insurance will be more likely to make claims than the population as a whole. Private health

insurance markets are possible only if either (a) health insurance is universal, so there is no selection and hence no adverse selection, (b) health insurance policies cover everyone in largely healthy sub-populations such as the employees of a large company, or (c) insurance companies issuing policies to individual purchasers are able to refuse insurance to individuals who look to be bad risks. A considerable portion of the population look to be bad risks because of their circumstances or health histories. They cannot be insured by voluntary private insurance.

The problems with adverse selection are among the perverse results of the asymmetries of information that characterize health care. In an ideal market, purchasers have the same knowledge of the characteristics of the commodities and services offered by various sellers that the sellers have, and sellers know everything about the preferences and purchasing power of the purchasers that the purchasers know. But health care and health insurance are not like that. Individuals may withhold information about their health from the providers and, more commonly, from insurers. Doctors frequently know more about the needs of patients and the efficacy and side-effects of possible treatments than patients; and they typically know more about patient's needs and about what alternatives are feasible in a particular context than do insurers.

For example, when my wife was in labor with our first child, the obstetrician (whom we were meeting for the first time) discovered that our son was breech—that is, he was on his way out butt first rather than head first. The obstetrician asked us, “Should I do a caesarean section, or should we proceed with a vaginal delivery?” How were my wife and I supposed to decide? Was I supposed to run off to the library, do research, and then discuss the matter with my wife (who, in the midst of labor was not in exactly ideal deliberative circumstances)? The informational asymmetry was extreme, and all that we could rationally do (apart from a few brief queries to the doctor) was to rely on the obstetrician's advice. There was no realistic possibility of shopping for a better price or a preferred provider, if indeed we had any grounds at that moment for preferring one obstetrician to another. We had to trust, and we relied on aspects of our medical system such as educational requirements, licensing, and codes of professional ethics to limit the risks of such trust. There is a big difference between employing someone to deliver a child and employing someone to repair one's shoes, and there is no reason to suppose that unregulated markets will work well in obstetrics because they work well in shoe repair.

An unregulated market in health care will not be efficient. A large portion of the population will be uninsured because they cannot afford comprehensive insurance or because of pre-existing conditions. As a result, their expensive needs will be unmet, which will be worse for them, for providers, and for productivity and economic growth. Providers will be able to exploit their superior knowledge concerning how to treat their patient's medical needs, subjecting those patients who are able to pay to unnecessary tests and treatments. Patients will find it hard to distinguish competent from incompetent providers. To remain in business, insurance companies will impose myriads of detailed provisions in insurance policies specifying precisely what they will and won't cover. To make their refusal to cover pre-existing conditions effective, insurance companies will devote considerable resources to denying benefits to individuals who may have concealed

details of their health histories. Without universal programs such as Medicare and Medicaid, mandatory emergency care, and regulations on health care providers such as licensing and codes of ethics, Senator Johnson's simple market would be a nightmare exaggeration of the chaotic and outlandishly expensive US health care system as it was before that Affordable Care Act. As I will argue in more detail in Section 23.5, an efficient health insurance system requires either coordinated centralized control by the government as the sole or ultimate insurer or very highly regulated private insurance markets like those in the Netherlands or Switzerland.

23.3 ETHICAL LIMITS ON EFFICIENCY: FAIRNESS, SOLIDARITY, AND COMPASSION

Readers may be impatient with the preoccupation thus far with efficiency. Whether or not a market system is efficient, many would object that the health care one receives ought not to depend on one's wealth. It strikes many as morally outrageous that one infant dies and another lives because the parents of the first are uninsured and cannot pay for a life-saving treatment. No one denies that health care should be efficient; but it must also be *fair*.

Fairness and efficiency do not always go hand-in-hand. Even if everyone is insured, allocating health-related resources by cost-effectiveness may be unfair as well as otherwise ethically objectionable. Although the many ethical objections to relying on cost-effectiveness are often described as resting on the value of fairness (for example, Daniels 1994 and Brock 2003), in my view some are matters of fairness and others are not. I think that in addition to efficiency there are at least four important values to which a health care system should be sensitive: fairness, compassion, solidarity, and choice (or freedom).

In this section, I shall discuss briefly three ethical objections that purported rest on concerns of fairness.⁶ First, if some medication, which is in short supply, is cheaper or more efficacious in treating one group of people than another, then it will be more cost effective to treat everyone in the first group before treating anyone in the second group. As population surveys show, this strikes most people as unfair. Call this this "the fair chances criticism."

Second, the pursuit of efficiency (as cost-effectiveness) implies that it makes no difference whether the health system provides a health improvement of a given magnitude to

⁶ For reasons of space I shall not discuss a fourth objection, which rests on qualms about aggregation. If there is a choice between providing a major benefit, such as preventing death to a small number of individuals and for the same cost providing a small benefit to a large number of people, then, if the number of people who get the small benefit is large enough, cost-effectiveness prescribes providing the small benefit to the large number. This strikes many people as morally objectionable. In their view, there is no number of headache cures that should take priority over saving someone's life. See Kamm (2002), Scanlon (1998), Kelleher (2014), and Voorhoeve (2014, 2017). I believe that this objection is misconceived, but the issues are intricate and require lengthy discussion.

someone who is moderately ill or to someone who is severely ill. Population surveys show that most people believe that the benefit should go to the individual who is more severely ill. Call this “the severity objection” (Nord 1999, 2013).

Finally, saving the lives of those who are free of disabilities results in a healthier population than saving the lives of those who are disabled. If the “effectiveness” in cost-effectiveness is effectiveness at improving population health and the costs are the same, then cost effectiveness says to save those without disabilities. This implication is obviously ethically repugnant. Call this “the discrimination objection.”

In my view, the discrimination objection really is a fairness objection, but the fair chances objection relies on the value of solidarity, while the severity objection derives from the value of compassion. To defend these claims requires saying something about fairness. In my view, Kant expressed the core idea of fairness in his second formulation of the categorical imperative, which insists that we treat rational agents as ends in themselves and never merely as means. I interpret those remarks as requiring that we show a particular sort of respect to one another (Darwall 2004). Unlike the sort of respect that is earned or lost by people’s choices, this respect involves an acknowledgement of one another as possessing a moral status that authorizes each of us to make claims of one another and forbids the sacrifice of important interests of one individual in order to provide equivalent or lesser benefits to others. Treating those who are disabled as having lesser claims to life saving is unfair in its failure to give an equal weight to their important interests.

Fairness is not exclusively an *ex ante* consideration governing how we treat people and hence affect their expectations. It is also a concern about outcomes. Allowing significant unchosen and avoidable inequalities in outcomes acquiesces in processes that, given their consequences, cannot be fair.

Although the fair chances objection at first glance appears to derive from concerns about fairness, I believe that this is a mistake. Suppose that one adopts a Rawlsian interpretation of the Kantian maxim (Rawls 1971). Behind the veil of ignorance, we reject as unfair putative principles of justice that sacrifice the fundamental interests of some in order to enhance the prospects of others. It is for this reason that Rawls argues against utilitarianism and in support of his difference principle. Even though the difference principle countenances inequalities, defenders of those inequalities can say to those who are least favored, “There is no way to arrange the basic structure of society so that the least favored, whoever they are, would be more advantaged than you are.” The favoring of those who can benefit more, to which the fair chances objection is directed, conforms to this interpretation of Kantian respect. One can truthfully say to those who are in the group for whom the treatment is less efficacious or more expensive, “There is no alternative to the cost-effective allocation of treatment in which those who are worst off would be doing better than you are or in which fewer individuals would be worse off.”

If one grants that the fair-chances objection does not point to any unfairness, it does not follow that there is nothing morally objectionable about treating all those who are cheaper to treat or who can benefit more from the treatment before treating anyone from other groups. One reason to object is if the members of the groups that are treated

last or not at all are disadvantaged in other ways. One reason why people may get less benefit from a treatment or why it may be more expensive to treat them may be that they are disabled or belong to disadvantaged minorities. In that case, the fair-chances objection fades into the discrimination objection.

When treating members of one group ahead of members of another does not involve unjust discrimination, I think it is still morally objectionable owing to its violation of solidarity. Solidarity is a tricky virtue (Prainsack and Buyx 2011), and when it contributes to suspicion and hatred of outsiders, it is more a vice than a virtue. But a recognition of our common humanity and vulnerability and the more detailed commonalities among those who share closer social bonds is crucial to political stability and flourishing community life. Solidarity is related to fairness because it is a recognition of something we share and of the equality or comparability of our circumstances. But rather than invoking our universal common rationality and moral capacities, solidarity focuses on our common vulnerabilities and aspirations at a more local level. Telling those who are more expensive to treat or who will benefit less from treatment to wait until all others have been treated is collectively turning our backs on fellow humans, citizens, neighbors or family members. It is objectionable because it displays the lack of a common bond rather than because it is unfair.

I am also unconvinced of the severity objection, which maintains that a benefit of a given magnitude should be given to those who are more severely ill before going to those who are less severely ill. The fact that one individual is more severely ill than another is typically a reason to tend to that individual first, because it is often possible to provide a larger benefit to those who are more severely ill, and because their condition, unlike the condition of someone who is only moderately ill, often will have irreversible consequences if treatment is delayed. But these reasons to attend to those who are more severely ill in preference to those who are less severely ill are not relevant to the severity objection. It maintains that it is unfair to a health benefit of exactly the same magnitude with just the same consequences to someone who is moderately ill when one might have given a benefit of the same size to someone who is more seriously ill. The severity objection assumes when the same benefit is to be had whether Ms. S (who is severely ill) or Mr. M (who is moderately ill) is treated, Ms. S should be treated.

A defender of the severity objection might invoke egalitarian or prioritarian (Parfit 1991) grounds. If one treats Ms. S rather than Mr. M, then their health states after the treatment will be more similar to one another, whereas if one treats Mr. M, one makes the existing inequality worse. But there is no good reason for an egalitarian to focus narrowly on inequalities in health. Ms. S may in other ways be far better off than Mr. M. Similar problems plague any attempt to defend priority for Ms. S on the grounds that the health care system should weight the benefits and burdens imposed on those who are worse off more heavily than the benefits and burdens of those who are better off. The fact that Ms. S is more severely ill does not imply that she is worse off. Moreover, the egalitarian and the prioritarian premises in these arguments are themselves in need of defense. The correlation between those who are most severely ill at a given moment and those who have been overall worst off over their lives is very weak.

I hypothesize that the severity objection is in part confusion and in part a reflection of compassion. The possibilities for confusion are multiple. Those surveyed may think that those who are most ill at the moment are worse off in general, or they may mistakenly assume that the same benefit when given to someone who is more severely ill makes a greater contribution (and is hence not the same benefit after all). Setting aside possible confusions, I think that the severity objection taps into a further value, compassion. Unlike impartial benevolence, which supports efficiency and cost-effectiveness, compassion is focused on individuals in unfortunate circumstances. Compassion does not permit us to turn one's back on those who are suffering and who are particularly vulnerable and at risk, even if denying treatment serves the common good, and is thus favored by impartial benevolence. Compassion is the sentiment that leads to wasteful treatment of those whose prospects for benefit are slim. We may all benefit from a rational determination that it is time to let Grandma go, but a society in which we gladly turn away, congratulating ourselves on our rational benevolence, could turn into a nightmare. If we have understood that treating Ms. S does not bring about a greater benefit and that she is not necessary worse off, all things considered than Mr. M, we may still favor treating her, because of our compassionate concern for her.

As I have just argued, the objections to allocating health-related resources in terms of cost effectiveness reflect the importance of three values: fairness, solidarity, and compassion, not just one. Whether I am right or wrong—that is, whether to assess health care policies in terms of efficiency and fairness or in terms of efficiency, fairness, solidarity, and compassion—the health care system faces the problem of how to make trade-off among different values. How much efficiency should it trade for how much greater fairness? How much efficiency or fairness should it sacrifice in order not to abandon those who are in desperate need?

Economic analysis cannot tell us how to make these tradeoffs, but it has a crucial role in revealing what the tradeoffs truly involve. Knowing that an unregulated market will be inefficient as well unfair is a crucial first step. Interventions in the market to make its outcome fairer need not diminish efficiency. But recognizing the inefficiency of unregulated health care markets is only a first step. Figuring out how to make the health care system efficient while at the same time heeding the other values is an enormously complicated task. For example, if providers are paid a fee for services, then they may provide more services than is cost effective or indeed good for their patients.⁷ Given the asymmetries in information, this kind of exploitation will be hard to root out. In contrast, paying providers a capitation fee may give them an incentive to provide fewer services than are needed. This is but one problem of many. There is a great deal of work for social scientists and philosophers to do in devising health care institutions that cope with the pathologies of insurance and the asymmetries of information in ethically acceptable ways.

⁷ In his examination of why the costs of treatment in McAllen, Texas, are so much higher than elsewhere, Atul Gawande (2009) finds that physicians there own laboratories and facilities and thus profit from tests and referrals.

23.4 FREEDOM AND CHOICE IN HEALTH CARE

I have left out one crucial value or family of values, which greatly complicates matters, while at the same time pointing, perhaps, to a way to reconcile the many values that are relevant to health care. That crucial value or family of values is freedom. I say that it may be a family of values because freedom encompasses (a) *negative freedom*, the mere absence of legal or social constraints; (b) *opportunity or choice*, the range of alternatives that are open to an individual; and (c) *autonomy or self-determination*, the extent to which the individual determines the course of his or her own life.⁸ Arguably, an unregulated market best protects negative freedom, but an unregulated market greatly limits the range of alternatives that are available to most people and the extent to which they can determine their own lives.

Freedom is crucial in thinking about health, because as Norman Daniels (1985, 2007) in particular has argued, the importance of health lies mainly in its bearing on freedom, in the sense of the range of alternatives that is open to an individual. Indeed, one might argue that, apart from the suffering that diminished health may entail, the importance of health lies entirely in its limitations on choice. On this view, one should measure the value of health by its bearing on opportunity rather than by its bearing on well-being. Whether or not the loss of a leg diminishes one's well-being, it diminishes the range of activities that are open to an individual. In addition, cognitive deficiencies and emotional disturbances may diminish autonomy.

But what is germane to this chapter is the importance of choice to health care, rather than the importance of health to choice (Engelhardt 1997; Lomasky 1981). Health care decisions are seldom purely technical, even if they are frequently guided by the advice of professionals who typically understand pathologies and their treatments better than patients do. Different treatments have different risks, and individuals may reasonably have different attitudes toward these risks. Treatments have different side effects in the short run and may lead to different long-run impairments. The importance of these differ across individuals. Moreover, health care decisions express attitudes toward the human body and mind and toward the significance of human lives. Witness, for example, the legal challenges in the United States over the Obamacare requirement that health plans cover the expense of contraceptives. Because of the importance of choice and self-determination, health care decisions should be responsive to individual values. Autonomous choice of treatment, to the extent that is feasible, is a crucial desideratum for health care that may compete with efficiency, fairness, solidarity, and compassion. Mandatory vaccination against whooping cough is cost effective, but it limits the choices of parents (as I am inclined to think it should). The conflict here between freedom and efficiency and the need to make a trade-off among different values are evident.

⁸ See Berlin (1969), MacCallum (1967), and Hausman et al. (2017: ch. 10).

How then should a health care system be devised that makes the best trade-offs among the five values that should govern health care policies? All of these values are important. Every acceptable health care system should place some weight on every one of these values. In much of the literature, efficiency has been the starting point or default, and authors have seen the problem as qualifying the use of cost-effectiveness information in order to accommodate other values, which are usually jumbled together as matters of fairness. But one might start with any one of the five values and then consider how to incorporate the others.

Let us begin by envisioning a health care system that is preoccupied solely with negative freedom—that is, the absence of legal or social constraints on individual choice. Such a health care system would not tax individuals (which limits their negative freedom) in order to provide health care for others, and unless private charity stepped in, it would not even cover emergency treatment. There might be a case for the treatment and prevention of epidemics, on analogy to the provision of national defense (Lomasky 1981), but otherwise negative freedom implies a completely “free” market in health care. In this market, there will still be a demand for the pooling of risks that health insurance provides, and it would be profitable for private insurance companies to form and sell risk-adjusted policies to individuals and policies to employees of companies adjusted to the company’s average risks. Given the asymmetry of information, there would be a demand for the certification of the quality of hospitals and caregivers. Presumably private firms would form to meet this demand.

Because the available resources are not sufficient to enable everyone to have access to every health care service that is likely to be beneficial, health care must be rationed: some people must be unable to access health care that could benefit them. In an unregulated market, this rationing would be done by the choices of individuals first among insurance policies (as constrained by their incomes and the cost of insurance) and then among health care services, as constrained by what their insurance will cover and the out-of-pocket expenses that individuals can pay for. The system will rely on market forces, including scrutiny and bargaining by individuals and insurance companies, to control the fees providers demand.

Although an unregulated market protects negative freedom, it fails miserably to satisfy the other values. The limited access to care it would provide to those who are poor or have pre-existing conditions would be unfair and inefficient. Efficiency would also be diminished by the administrative overhead implicit in private risk-adjusted health insurance. An unregulated private insurance market turns its back on individuals who could easily have been helped if they had been insured. Compassionate individuals condemn the deaths and suffering that could be averted by the sort of limits on negative freedom that taxation or insurance mandates would require. Also, to the extent to which market provision of health care tolerates such abandonment of members of one’s community, it undermines solidarity.

Furthermore, an unregulated market is extremely undesirable from the perspective of freedom itself, once one considers choice and self-determination as well as the absence of constraint. A concern with freedom rather than well-being does not make the difficulties of asymmetric information, or the problems caused by health insurance,

disappear. The ability to choose requires protection from predatory health care providers who exploit their superior knowledge of patient's needs. Effective individual choice requires reliable sources of information and advice, which in turn may require licensing and educational requirements on providers. Defenders of markets argue that market forces would reduce costs, so that few would be uninsured because of an inability to pay for it. However, because of asymmetric information and the difficulties of "exit" and of shopping around for care, it is questionable whether largely unregulated markets will be able to control costs. The experience in the United States is not encouraging, but, in fairness, one must note that the modern United States has never had a purely market-driven health care system. Quite apart from the consequences for efficiency, a failure to control costs limits the range of services that individuals will be able to access and hence limits opportunities.

An unregulated private health insurance system comes with serious limitations on opportunity and self-determination. Unregulated health care and health insurance markets leave both the poor and (owing to adverse selection) those with pre-existing conditions without health insurance. The unavailability of care for a large portion of the population severely diminishes what people can choose. Furthermore, unregulated private health insurance limits the choices of many of those who are insured. Individuals who would have trouble purchasing health care on the open market may be locked into their jobs or even into marriages.

These are serious problems, even if one focuses exclusively on freedom; and they are not easy to cure. Community-rated insurance with guaranteed issue is not possible without limiting people's choices and forcing most people to purchase health insurance that meets minimum standards. Subsidizing health insurance for the poor and for those with pre-existing conditions requires substantial taxes, which necessarily limit other choices. There is no health care institution that does not limit freedom. Respecting freedom either means that the impoverished and uninsured patient who shows up at the emergency room with a heart attack will be tossed out into the street (unless some benefactor can be immediately located), or it means that everyone will either have to pay taxes to cover this care or will to pay more for health care to cover emergency care for the otherwise uninsured. However the trade-offs among different values are finally made, there is no system that promotes freedom in all of its senses. Completely protecting negative freedom limits choices and self-determination. Protecting choices and self-determination requires limiting negative freedom.

23.5 FIXING HEALTH CARE AND HEALTH INSURANCE MARKETS

What should be done to meet the demands of compassion, solidarity, fairness, and efficiency to a greater extent while compromising among the different aspects of freedom? Working out the details is a task for a lifetime or for a small army of economists

and philosophers, and it is impossible in what remains of this chapter. Moreover, there is no reason to suppose that the best compromise for one society will be best for others. All I can do here is to lay out the most important choices that have to be made.

Let us start with efficiency. Unregulated health care markets are inefficient both with respect to the promotion of population health and, more generally, with respect to economic productivity and growth. They are ineffective at controlling costs because of incentives to overtreat (and the asymmetric information that make overtreatment easy to get away with) and because of the huge administrative overhead implicit in private insurance. In addition, such markets limit access to health care to the poor and to those with pre-existing conditions. Doing so reduces population health and economic productivity.

There are two cures. The first is radical and steps away from markets, thereby limiting negative freedom as well as choice and self-determination, and eliminates private insurance altogether. This path substitutes some agency of government as a single insurer providing community-rated universal health insurance paid for via taxation. The universality eliminates the problem of adverse selection and the administrative effort devoted to identifying those with pre-existing conditions and denying their claims. Gone also are the marketing costs of private insurance and the burden on both providers and patients of dealing with the complicated policies of multiple insurance companies. Providers may be employees of the government agency responsible for health care or independent contractors paid for their services. A single payer system need not eliminate all choices, but minimally it eliminates choice of basic insurance. It may permit additional private insurance, or it may forbid it. As a glance around the world reveals, there are many ways to design a single-payer universal health care system.

A single-payer system is not necessarily the best cure for the inefficiencies of basic market health care, because without the existence of competition, government control may become sclerotic or even corrupt. An alternative is to harness private insurance companies to provide universal health care. This can be done by requiring that everyone purchase community-rated health care that meets minimum standards (so as to eliminate adverse selection) and requiring, as is the case in the Netherlands, that all insurance companies offer the basic package at the same price. Private insurance companies can then compete for customers by contracting with providers to provide better care or by offering additional insurance. Allowing private insurance companies to compete for the provision of basic health care increases choice only slightly because the basic insurance plan is uniform and mandatory. Because everyone will be required to purchase insurance, those who cannot afford to do so must be subsidized via a system of health insurance vouchers or in some other way. There are many detailed ways to structure a regulated private insurance system, just as there are many forms of single-payer systems, and it is hard to tell which is most efficient. In addition, it may be (as is perhaps the case with the health care arrangements in Singapore) that systems that are efficient in some geographical and cultural settings will be inefficient or even impossible in others.

Regardless of the institutional form, the provision of health care will be efficient to the extent to which it is cost effective. With the exception of inexpensive items like aspirin or

band-aids, which are more efficiently paid for out of pocket, the basic universal health insurance package will cover a set of health services that are all at least as cost-effective as those health services that are not covered. If health insurance purchasers are rational and well-informed, they will demand—and responsible insurance companies will offer—supplemental insurance plans that are cost effective. Concerns about efficiency severely qualify the reliance on markets.

What about fairness? Fairness has two faces. *Ex ante*, fairness in health care is concerned with the distribution of risks of illness, the availability of insurance, and the respect implicit in the procedures health care institutions employ. *Ex post*, fairness depends on how health and wealth outcomes are distributed. Enormous arbitrariness, both biological and social, is overlaid on the pattern of health outcomes determined by individual choices. Most health deficiencies arise either from causes that are unfair, like the lead poisoning of children who grow up in dangerous housing, or from chance factors that are morally arbitrary. One child grows up with abusive and neglectful parents; another has a loving home. One of two children from similar homes has leukemia; the other is healthy. The health care and public health system can lessen these inequalities, but it cannot eliminate them.

Health outcomes are not fair. Although there are inequalities that a better health care system can address, there are also inequalities that no health care system is capable of eliminating. As the result of chance or nature, one might hesitate to pass moral judgment on them, but whether one chooses to call them unfair or neither fair nor unfair, health outcomes are not distributed according to any theory of fairness. It might be possible to make the distribution of health fairer by diminishing the health of the healthy, but I shall suppose that such leveling down is ruled out by efficiency, compassion, and solidarity.

In addition to the arbitrary inequalities in health outcomes that health care is incapable of remedying, coverage limits within health insurance mean that among health conditions that are equally debilitating or risky, some will be treated and some will not be treated. If the treatment of condition A is less expensive than the treatment of condition B, then even if those with condition B are just as badly off as those with condition A, their treatment may not be cost effective. Although it may be fair *ex ante* to adopt a policy of treating A before turning to B, the resulting outcome is not fair. Efficiency and fairness are not orthogonal, but they do not walk side-by-side either. To be fair, outcomes should not be arbitrary. They should reflect individual choices, but health care cannot accomplish this, and it may create new inequalities. The problems of discrimination against the disabled and others who are badly off are instances of these difficulties.

What should be done? How should the conflict between efficiency and fairness be managed? What is the best trade off? Universal health insurance guarantees *ex ante* fairness with respect to access to everything covered by the basic health insurance package. That goes a long way toward satisfying the demands of *ex ante* fairness. Some health care systems go further, prohibiting supplemental private insurance, but doing so will not make the distribution of health expectations fair. Universal health care can be justified in terms of efficiency and, more powerfully, in terms of opportunity. The problems of discrimination against the disabled and against those who are more expensive to treat

largely dissolve if basic health care specifies which treatments it will and will not cover without regard to unrelated disabilities or non-health contingencies that make treatment more difficult or less efficacious. Conflicts remain but the specification of general rubrics for care rules out the most egregious discrimination. The inefficiencies these give rise to are in practice minor, because it is difficult and expensive to condition coverage limitations on details of the recipients' circumstances.

With respect to *ex post* fairness: it is not to be had this side of the grave. There is no distribution of health care that makes the outcomes fair. Universal insurance may lessen those inequalities in outcomes that are due to inequalities in treatment, and it typically (although not necessarily) lessens the inequalities in overall health outcomes. But even if everyone has access to the same excellent treatment, outcomes will differ.

Finally come the problems of making the health care system consistent with the demands of compassion and solidarity. I believe that compassion is impossibly demanding. With limited resources, there must come a time when the health care system, by some institutional means or another, says no. In poor countries the health authority or the impersonal market must say no to many efficacious and reasonably inexpensive treatments. In rich countries, the health care system does not have to say no nearly so often, but health insurance cannot cover extremely expensive treatments of limited or uncertain value. Inevitably, people in desperate need will be deprived of potentially beneficial treatment, and compassionate individuals will cry out in protest against cold-hearted "rational" health care policy.

Compassion's demand not to abandon those in need cannot be satisfied. Although any given no could be a yes, there must be no's—and lots of them. The best that any health care system can do to accommodate the demands of compassion is first of all to conform (to a reasonable degree of approximation) to the requirements of efficiency and fairness discussed above, so that those who are refused treatment are not refused treatment unnecessarily or unfairly. Beyond that, the health care system should recognize and encourage compassion by instituting an appeals procedure whereby those who object to denying treatment to someone can be heard. In addition, the health care should set aside some small allotment of resources to provide some compassionate but cost-ineffective care. Private charity can also help to soften the objectionable apparent cold heartedness of the health care system. This is largely a symbolic gesture, but an important one. Even as the health care system must reject the demands of compassion, it should recognize their worth. Finally, the health care system should justify the deprivation of care cogently by the larger benefits that the savings in resources provides to others.

I have the least to say about solidarity. As mentioned at the beginning of this chapter, how the health care system treats those who are suffering, vulnerable, and frightened is not just a matter of individual compassion but a matter of how we as members of some group—whether it be humanity as a whole or, more saliently, smaller more substantive groups—care for one another. Solidarity is relevant to the design of health care institutions because it bears on how one conceives of health insurance. Is health insurance a social device whereby we spread the risks of disease and disability or is it a device enabling

each individual separately to promote his or her own health more efficiently? From the perspective of solidarity, one should condemn risk-adjusted health insurance whereby those who need health care the most face the highest premiums. Their health problems are already a burden that their fellow citizens should seek to lighten rather than to aggravate. From an opposing perspective that rejects solidarity, those who are fortunate owe nothing to those who are unlucky, and it is unjust to make the fortunate pay more than the actuarial value of the risks they face.

If individuals have a choice of buying insurance beyond the required coverage, complaints about refusal to cover treatments can sometimes be answered by pointing out that individuals could have purchased insurance that would have covered that condition. In that case the responsibility for the denial of coverage is at least in part the individual's own. That helps to legitimate the denial, but it also threatens solidarity by making protection against disease and injury an individual choice rather than a social obligation.

23.6 CONCLUSION

In casual remarks on February 28, 2017, President Donald J. Trump remarked, "Nobody knew health care could be so complicated." He probably spoke sincerely, and he was right that health care is complicated. But he was of course wrong to suppose that nobody knew this. This chapter has argued that these complications are both ethical and economic, and that both moral reflection and economic analysis are essential to understanding and addressing them. At least five different values are relevant: efficiency, fairness, freedom, compassion, and solidarity. Each of these is in turn a complicated notion in need of more elaborate and precise development than I have been able to provide here. The economic complications are just as great as the philosophical. Markets for health care differ dramatically from markets for services such as hair cutting or lawn mowing. Needs for health care services are intermittent, often unpredictable, and sometimes both urgent and extremely expensive, and possession of knowledge about health care needs and how to treat them is asymmetric. Insurance is necessary to cope, but it brings in its wake a host of further complications. Health care is extremely complicated. Designing a health care system requires joint effort by philosophers, economists, and others to find a way through the complications.

ACKNOWLEDGMENTS

I am grateful for the comments of an excellent audience at the Erasmus Institute for Philosophy and Economics. This chapter was completed while I was a Ludwig Lachmann Fellow at the London School of Economics, and I am grateful to the support and hospitality of the Department of Philosophy, Logic and Scientific Method and the Centre for the Philosophy of the Natural and Social Sciences.

REFERENCES

- Arrow, Kenneth. 1963. "Uncertainty and the Welfare Economics of Medical Care." *American Economic Review* 53: 941–73.
- Berlin, Isaiah. 1969. "Two Concepts of Liberty." In *Four Essays on Liberty* (Oxford: Oxford University Press), pp. 118–72.
- Boorse, Christopher. 1977. "Health as a Theoretical Concept." *Philosophy of Science* 44: 542–73.
- Boorse, Christopher. 1997. "A Rebuttal on Health." In James M. Humber and Robert F. Almeder (eds.), *What is Disease?* (Totowa, NJ: Humana Press), pp. 1–134.
- Boorse, Christopher. 2011. "Concepts of Health." In Fred Gifford (ed.), *Philosophy of Medicine: Handbook of Philosophy of Science, Volume 16* (Amsterdam: Elsevier), pp. 13–64.
- Brock, Dan. 2003. "Ethical Issues in the Use of Cost-Effectiveness Analysis for the Prioritization of Health Care Resources." In Tessa Tan-Torres Edejer et al (eds.), *Making Choices in Health: WHO Guide to Cost-Effectiveness Analysis* (Geneva: World Health Organization), pp. 289–311.
- Broome, John. 2002. "Measuring the Burden of Disease by Aggregating Well-Being." In Christopher Murray et al (eds.), *Summary Measures of Population Health: Concepts, Ethics, Measurement and Applications* (Geneva: World Health Organization), pp. 91–113.
- Daniels, Norman. 1985. *Just Health Care*. Cambridge: Cambridge University Press.
- Daniels, Norman. 1994. "Four Unsolved Rationing Problems." *Hastings Center Report* 24 (4): 27–9.
- Daniels, Norman. 2007. *Just Health*. Cambridge: Cambridge University Press.
- Darwall, Stephen. 2004. "Respect and the Second-Person Standpoint." *Proceedings and Address of the American Philosophical Association* 78: 43–55.
- Dolan, Paul. 1997. "Modeling Valuations for Euroqol Health States." *Medical Care* 35: 1095–108.
- Dolan, Paul, and Daniel Kahneman. 2008. "Interpretations of Utility and their Implications for the Valuation of Health." *Economic Journal* 118: 215–34.
- Engelhardt, H. Tristram, Jr. 1974. "The Disease of Masturbation: Values and the Concept of Disease." *Bulletin of the History of Medicine* 48: 234–48.
- Engelhardt, H. Tristram, Jr. 1997. "Freedom and Moral Diversity: The Moral Failures of Health Care in the Welfare State." *Social Philosophy and Policy* 24: 180–96.
- Finkelstein, Amy, with Kenneth J. Arrow, Jonathan Gruber, Joseph P. Newhouse, and Joseph E. Stiglitz. 2015. *Moral Hazard in Health Insurance*. New York: Columbia University Press.
- Gawande, Atul. 2009. "The Cost Conundrum: What a Texas Town Can Teach us about Health Care." *The New Yorker*, June 1. Available at <https://www.newyorker.com/magazine/2009/06/01/the-cost-conundrum>.
- Griffin, James. 1986. *Well-Being: Its Meaning, Measurement and Moral Importance*. Oxford: Clarendon Press.
- Hausman, Daniel. 2012a. "Health, Naturalism, and Functional Efficiency." *Philosophy of Science* 79: 519–41.
- Hausman, Daniel. 2012b. *Preference, Value, Choice, and Welfare*. Cambridge: Cambridge University Press.
- Hausman, Daniel. 2015. *Valuing Health: Well-Being, Freedom, and Suffering*. New York: Oxford University Press.
- Hausman, Daniel, Michael McPherson, and Debra Satz. 2017. *Economic Analysis, Moral Philosophy, and Public Policy*, 3rd edn. New York: Cambridge University Press.

- Johnson, Ron. 2017. "Where the Senate Health Care Bill Fails." *New York Times*, June 26. Available at <https://www.nytimes.com/2017/06/26/opinion/senate-health-care-bill.html>.
- Kahneman, Daniel. 1999. "Objective Happiness." In Daniel Kahneman, Ed Diener, and Norbert Schwarz (eds.), *Well-Being: Foundations of Hedonic Psychology* (New York: Russell Sage Foundation Press), pp. 3–27.
- Kahneman, Daniel. 2000a. "Evaluation by Moments: Past and Future." In Daniel Kahneman and Amos Tversky (eds.), *Choices, Values and Frames* (New York: Cambridge University Press and Russell Sage Foundation), pp. 693–708.
- Kahneman, Daniel. 2000b. "Experienced Utility and Objective Happiness: A Moment-based Approach." In Daniel Kahneman and Amos Tversky (eds.), *Choices, Values and Frames* (New York: Cambridge University Press and Russell Sage Foundation), pp. 673–92.
- Kahneman, Daniel, and Alan Krueger. 2006. "Developments in the Measurement of Subjective Well-being." *Journal of Economic Perspectives* 20: 3–24.
- Kahneman, Daniel, and Robert Sugden. 2005. "Experienced Utility as a Standard of Policy Evaluation." *Environmental and Resource Economics* 32: 161–81.
- Kahneman, Daniel, and Richard Thaler. 2006. "Utility Maximization and Experienced Utility." *Journal of Economic Perspectives* 20: 221–34.
- Kamm, Frances. 2002. "Health and Equity." In Christopher Murray et al. (eds.), *Summary Measures of Population Health: Concepts, Ethics, Measurement and Applications* (Geneva: World Health Organization), pp. 685–706.
- Kelleher, J. Paul. 2014. "Relevance and Non-Consequentialist Aggregation." *Utilitas* 26: 385–408.
- Lane, Harlan. 2002. "Do Deaf People Have a Disability?" *Sign Language Studies* 2: 356–79.
- Lomasky, Loren. 1981. "Medical Progress and National Health Care." *Philosophy & Public Affairs* 10: 65–88.
- MacCallum, Gerard. 1967. "Negative and Positive Freedom." *Philosophical Review* 76: 312–34.
- Menzel, Paul. 2012. "Justice and Fairness: A Critical Element in U.S. Health System Reform." *Journal of Law, Medicine, and Ethics* 40: 582–97.
- Murray, Christopher. 1996. "Rethinking DALYs." In Christopher Murray and Alan Lopez (eds.), *The Global Burden of Disease: A Comprehensive Assessment of Mortality and Disability from Diseases, Injuries, and Risk Factors in 1990 and Projected to 2020* (Boston, MA: Harvard School of Public Health), pp. 1–98.
- Murray, Christopher, Joshua Salomon, Colin Mathers, and Alan Lopez (eds.). 2002. *Summary Measures of Population Health: Concepts, Ethics, Measurement and Applications*. Geneva: World Health Organization.
- Murray, Christopher, et al. 2012. "The Global Burden of Disease Study 2010." *Lancet* 380: 2053–260.
- Nord, Erik. 1999. *Cost-Value Analysis in Health Care: Making Sense Out of QALYs*. Cambridge: Cambridge University Press.
- Nord, Erik. 2013. "Priority to the Worse Off: Severity of Current and Future Illness Versus Shortfall in Lifetime Health." In Nir Eyal et al. (eds.), *Inequalities in Health: Concepts, Measures, and Ethics* (Oxford: Oxford University Press), pp. 66–73.
- Nordenfelt, Lennart. 2000. *Action, Ability and Health: Essays in the Philosophy of Action and Welfare*. Dordrecht: Kluwer.
- Parfit, Derek. 1986. *Reasons and Persons*. Oxford: Oxford University Press.
- Parfit, Derek. 1991. *Equality or Priority*. The Lindley Lecture. Department of Philosophy, University of Kansas.

- Prainsack, Barbara, and Alena Buyx. 2011. *Solidarity: Reflections on an Emerging Concept in Bioethics*. Swindon: Nuffield Foundation.
- Rawls, John. 1971. *A Theory of Justice*. Cambridge, MA: Harvard University Press.
- Reznek, Lawrie. 1987. *The Nature of Disease*. London: Routledge & Kegan Paul.
- Scanlon, Thomas. 1998. *What We Owe to Each Other*. Cambridge, MA: Harvard University Press.
- Sumner, L.W. 1996. *Welfare, Happiness and Ethics*. Oxford: Clarendon Press.
- Voorhoeve, Alex. 2014. "How Should We Aggregate Competing Claims?" *Ethics* 125: 64–87.
- Voorhoeve, Alex. 2017. "Why One Should Count Only Claims with which One Can Sympathize." *Public Health Ethics* 10: 148–56.
- Wakefield, Jerome C. 1992. "The Concept of Mental Disorder: On the Boundary Between Biological Facts and Social Values." *American Psychologist* 47: 373–88.
- Wakefield, Jerome C. 1999. "Evolutionary Versus Prototype Analyses of the Concept of Disorder." *Journal of Abnormal Psychology* 108: 374–99.

CHAPTER 24

DEONTOLOGICAL MORALITY AND ECONOMIC ANALYSIS OF LAW

EYAL ZAMIR AND
BARAK MEDINA

24.1 INTRODUCTION

FOR several decades, economic analysis of law has been one of the leading—if not *the* leading—perspectives in legal theory, with immense contributions to legal analysis in practically every field of law (Posner 2014; Cooter and Ulen 2012). Cost-benefit analysis (CBA) forces one to consider the interrelationships between goals, means, incentives, and outcomes, in systematic and sophisticated ways. The very act of economic modeling compels one to determine the crucial variables pertinent to any issue. At the same time, even avid supporters of economic analysis can hardly deny the normative flaws that exist in standard economic analysis (Kornhauser 2000; Posner 2015). In particular, the criteria of economic efficiency tend to ignore fundamental ethical norms, such as the inherent immorality of deliberately harming other people. The consequentialist nature of economic analysis—namely, its denial of the intrinsic value of any factor other than the goodness of outcomes—makes it normatively unacceptable for many philosophers and lawyers. In fact, it troubles lawyer-economists as well (Ulen 2015; Calabresi 2016; Posner and Sunstein 2017).

Deontological moral theories maintain that the goodness of outcomes counts, but that it is not the only morally relevant factor (Rawls 1999: 26; Kagan 1998: 60, 64, 70–8; Kamm 1993: 76). The pursuit of good outcomes is subject to *constraints*. Certain acts are inherently wrong and therefore impermissible, even as a means to furthering the overall

good. The primary constraint is against harming other people. Additional constraints prohibit conduct such as lying and breaking promises. Currently prevailing deontological theories are *moderate* rather than *absolutist*. They admit that constraints have thresholds. A constraint may be overridden for the sake of furthering good outcomes, or avoiding bad ones, if sufficient good (or bad) is at stake (Thomson 1986; Brennan 1995; Kagan 1998: 78–80; Zamir and Medina 2010: 46, 51–6). Thus, while consequentialism (presumably at least) approves of the deliberate killing of one innocent person to save the lives of two others, moderate deontology might justify such killing only for the sake of saving a far greater number (perhaps hundreds or thousands) of people. Similarly, while consequentialism supports breaking a promise whenever it produces even a slightly greater net benefit than keeping one's word, moderate deontology justifies breaking a promise only to avoid very considerable losses (and an absolute deontologist would object to killing or breaking a promise under any circumstances). Deontology differs from consequentialism not only in holding that there are constraints to attaining the best outcomes, but also in its recognition of *options*: people are sometimes allowed to refrain from maximizing the good, even if no constraint infringement is involved.

This chapter distinguishes between positive and normative economic analysis, and briefly describes the tenets of normative economic analysis of law and the challenges it faces. It then focuses on the consequentialist nature of standard economic analysis and the deontological critiques thereof. It argues that ignoring deontological constraints and options is not a viable option for legal policymaking, and critically examines several methods of incorporating them into the analysis, their prospects and limitations. Such incorporation may arguably improve economic analysis of law, and economic analysis in general, as a normative theory and a framework for policymaking, and possibly as a descriptive and predictive tool as well.

24.2 NORMATIVE ECONOMIC ANALYSIS OF LAW

Economics is conventionally divided into *positive* and *normative* fields. Positive economic analysis explains and predicts human behavior—and social outcomes—based on rational choice theory, which assumes that people act as rational maximizers of their own utility. Normative (or welfare) economics—which is the focus of our study—evaluates the desirability of acts, rules, and so on according to their efficiency. Unlike typical economic analyses, most economic analyses of law are explicitly normative, as they refer to the optimal design of the law, which is a normative system. While efficiency analysis might focus on acts, projects, rules, policies, and so on, the focus of normative economic analysis of law is usually legal norms and institutions.

Normative economics is consequentialist—that is to say, it maintains that the desirability (efficiency) of anything depends solely on the goodness of its outcomes. More

specifically, it is a *welfarist* theory. It contends that the only factor that ultimately determines the goodness of outcomes is the effect on individuals' welfare. The theory of the good underlying normative economic analysis is *preference satisfaction*, according to which people's well-being is enhanced to the extent that their desires are fulfilled (rather than a hedonistic, or an objective-list theory of human well-being). It resembles utilitarianism in assigning equal weight to the welfare of every person, but differs from the latter in its theory of human welfare—preference satisfaction, rather than feelings of pleasure and pain. In taking into account the welfare of all people, and attributing equal weight to the welfare of each person, welfare economics is a *universal* consequentialist theory (in contradistinction, for example, to ethical egoism).

Welfare economics ordinarily uses one of two criteria of efficiency: *Pareto* or *Kaldor-Hicks*. According to the Pareto principle, state of the world A is socially preferable (or *Pareto superior*) to state B if at least one person prefers A to B and all others are either indifferent between the two states or prefer A to B. State A is a *Pareto optimum* if there is no other possible state that is socially preferable to A in the above sense. According to the Kaldor-Hicks criterion, a state A is socially preferable to state B if those who prefer A to B gain, in monetary terms, from being in A rather than B, more than those who prefer B to A lose. Thus, a social change that does not meet the Pareto criterion should still be implemented if the gainers from the change can compensate the losers and remain better off (Posner 2001: 95–141). As long as economic analysis pertains to people's voluntary behavior (such as market transactions), it may reasonably use the Pareto criterion, assuming that rational people would not consent to anything that makes them worse off. However, the Pareto criterion is hardly suitable for analyzing legal policies, because any conceivable legal rule will make some people worse off than some other conceivable legal rule, and thus no rule is Pareto superior to all others (Calabresi 1991; Dorff 2002). Even a purely facilitative legal rule—such as “contracts should be performed” (which leaves people free not to make contracts)—is not Pareto superior to all other rules. This is because some people might sensibly prefer that the resources the state uses to enforce contracts be used to pursue other goals, while others might be better off under a rule that states “contracts must be performed, unless one of the parties changes her mind within [a certain period of time].”

Each of these features of economic analysis is subject to criticism (Hausman et al. 2017). The very reliance on people's preference satisfaction as the yardstick for human welfare is problematic, because people's preferences are often influenced by cognitive biases and faulty reasoning (Keren and Wu 2015; Zamir and Teichman 2018); overly modest (Sen 1985); or objectionable, anti-social, or even abhorrent (Griffin 1986). More fundamentally, the theory of the good underlying economic analysis may be faulted for denying the intrinsic value of anything but human welfare, such as the natural environment or the well-being of animals (Degrazia 1996). Its disregard for notions such as *desert* and *fairness* is likewise objectionable (Kagan 1998: 54–9, 309). For example, standard economic analysis tends to object to laws that prohibit marketplace discrimination. Arguably, if discriminatory practices were inefficient, their victims would have “purchased” from the bigots their freedom to discriminate. Thus (according to this argument),

the endurance of discrimination indicates that the discriminators value their freedom to discriminate more than the victims value the entitlement of not being discriminated against. Therefore, anti-discrimination laws are inefficient (Posner 1987; Epstein 1992; Vassel 1994).

Further concerns are raised by the measurement of welfare in monetary terms, including the claim that some things are not commensurable with money (Chang 1997), and the regressive effect of *willingness to pay* (WTP) as a measure of people's utility. More fundamentally, in its basic form, the Kaldor-Hicks criterion only measures total welfare, and attributes no intrinsic value to its distribution among people. Each of these critiques have spurred extensive debates, many of which are discussed elsewhere in this handbook.¹ Our focus, however, is on the consequentialist nature of welfare economics, and the deontological critique leveled against it, which we will now discuss.

24.3 DEONTOLOGICAL CONSTRAINTS AND OPTIONS

24.3.1 General

Deontological moral theories are typically more complex than consequentialist theories, and more diverse. They do, however, share a common denominator: all deontological theories view the goodness of outcomes as a morally relevant factor, but, unlike consequentialism, they do not consider it the *only* intrinsically important factor (Rawls 1999: 26; Kamm 1993: 76). Deontological theories prioritize such values as autonomy, human dignity, basic liberties, truth-telling, fidelity, fair play, and keeping one's promises, over the promotion of good outcomes. They include constraints on attaining the best outcomes (Kagan 1998: 70–152). The primary deontological constraint is against harming other people. This constraint does not apply to each and every interest an individual may have. The precise list of deontological constraints is debatable, but it usually includes restrictions on violating fundamental rights, such as the rights to life and bodily integrity, human dignity, and freedoms of religion and speech. It also includes special obligations arising from promises and agreements, and restrictions on lying and betrayal (Kagan 1998: 84–9; Darwall 2003a). In addition, there is a “deontological requirement of fairness, of evenhandedness or equality in one's treatment of people” (Nagel 1986: 176; see also Kamm 1993).

At the same time, deontology holds that requiring people to impartially promote the overall good, irrespective of their own interests and the interests of their loved ones, runs counter to human nature and with the notion of people as separate entities. Therefore, at least under some circumstances, people may—or perhaps even should—prefer their own welfare, or the welfare of their family, friends, or community, over the overall good

¹ See, in particular, the chapters by Adler, DeMartino, White, and Wight.

(Nagel 1986: 164–207; Kagan 1998: 161–70; Scheffler 2001: 1–130; Mulgan 2001). Thus, deontological theories recognize agent-relative *constraints* (on promoting the good) and agent-relative *options* (not to promote the good).

While consequentialism judges the morality of an action (or anything else) according to its outcomes, deontology focuses on the morality of the action itself and on relations between people (Darwall 2003a). For example, a consequentialist would maintain that, all else being equal, the fact that one person is tortured is a bad thing—and if two people are tortured, it is certainly worse. In contrast, a deontologist would argue that it is immoral for an agent to torture someone else, or to be in a torturer–victim relationship. For the deontologist, the fact that she tortures one person may well be worse than the fact that somewhere in the world, two people are being tortured by someone else. As ordinarily conceived, therefore, consequentialism adopts an impartial or agent-neutral perspective, whereas deontology is distinctively agent-relative.

24.3.2 Doing and Allowing; Intending and Foreseeing

The notion of agent-relativity implies that there must be a difference between a person's duty to refrain from violating a constraint and her duty not to bring about, or to prevent, other violations—even where such violations are the expected outcome of avoiding the current one. Otherwise, the prohibition to kill one person in order to save two others would prohibit both killing the person and not killing her (thereby causing the death of the two). Deontology therefore resorts to a distinction between actively violating a constraint and not preventing constraint violations by others, or some such. In the context of the constraint against harming people, deontology therefore distinguishes between *actively harming* a person and *not aiding* her—often labeled the *doing/allowing* distinction (Steinbock and Norcross 1994; Howard-Snyder 2016). While doing harm is at least presumably immoral, allowing harm is not ordinarily so. This is not to say that the duty to help others never exists. Deontology does require agents to come to other people's aid—at least when the loss or suffering experienced by those others is sufficiently large, and the risks and costs to the agent do not exceed a certain threshold. What the doing/allowing distinction suggests is that the moral responsibility for actively harming people is greater than for failing to prevent harm, and therefore the former is worse than the latter (Scheffler 2004).

Another distinction deontologists often draw is between *intending* harm, and merely *foreseeing* it. Intending harm is immoral even if the harm is merely allowed, while foreseeing harm is not necessarily immoral (Nagel 1986: 179–80; Kagan 1989: 128–82; Bennett 1995: 194–225). The constraint against intending harm forbids not only harming a person as an end, but also as a means to attaining some other goal. Using a person as a means to an end violates the Kantian duty to respect people as ends in their own right. A related distinction is thus drawn between harming a person as a *side-effect* of aiding or saving other people, and using a person as a means to aiding or saving others (McIntyre 2001, 2014; Alexander 2016).

The means/side-effect distinction is often discussed in reference to the “trolley problem” (Foot 1978; Thomson 1985; Kamm 2016). Imagine that an uncontrolled trolley is hurtling down a track. Five people stand directly in its path, who cannot escape and will be killed by the runaway trolley. An agent can flip a switch, diverting the trolley to another track, where it would kill a single individual. Should the agent flip the switch? Alternatively, suppose that the only way the agent can save the five people is by pushing another individual onto the track, blocking the trolley and killing that individual. Should the agent push the other individual? Most people find diverting the trolley morally permissible—perhaps even obligatory—while pushing the individual morally repugnant (Mikhail 2011). Deontologists explain the difference between the two by the distinction between killing as a mere side-effect (in the diversion scenario) and killing as a means (in the pushing scenario).

The doing/allowing and the intending/foreseeing distinctions usually overlap, with the former distinction sometimes rationalized as a proxy for the latter. In most cases of merely allowing harm, the harm is an unintended consequence of one’s inaction, and so the two distinctions coincide. The two distinctions do, however, yield different conclusions when one intentionally allows harm (thus violating only the constraint against intending harm), and when one actively harms another as a mere side-effect of achieving another goal (thus violating only the constraint against doing harm). These distinctions—along with others that deontologists make—play a very minor role, if at all, in consequentialist analysis.

24.3.3 Absolutist and Moderate Deontology

Deontological moral theories are either *absolutist* or *moderate*. While absolutist deontology maintains that constraints must not be violated for any amount of good consequences, moderate deontology holds that constraints have *thresholds*. A constraint may be overridden for the sake of furthering good outcomes or avoiding bad ones if sufficient good or bad is at stake. For example, even the constraint against actively or intentionally killing an innocent person may be justifiably infringed if such killing is the only way to save the lives of thousands of people (Thomson 1986; Brennan 1995; Moore 1997: 719–24; Kagan 1998: 78–80; Zamir and Medina 2010). The thresholds that have to be met to justify the infringement of other constraints, such as those against lying or breaking one’s promise, are much lower. Correspondingly, deontological options need not be absolute: when enough (net) good or bad outcomes are at stake, there is no longer an option not to further the good or avoid the bad.

In determining the amount of good or bad outcomes that may justify infringing a constraint, a moderate deontologist might reasonably take into account the distinctions between doing and allowing and between intending and foreseeing. Thus, the threshold that has to be met to justify actively harming another person when the harm is intended is plausibly much higher than when it is a mere side-effect. In the case of the trolley problem, it stands to reason that diverting the uncontrolled trolley to another track (that

is, killing as a side-effect) may be justified for attaining a net saving of a small number of lives, whereas pushing a person onto the track to block the trolley (killing as a means), is only justified for the sake of saving very many people. Now suppose that a person who is stuck on the track is expected to block the runaway trolley, thereby saving the lives of other people on the track. Arguably, not helping that person get off the track, thus allowing her to be killed as a means to save the others, is permissible for the sake of saving fewer people than in the pushing scenario (which involves an active killing).

24.3.4 Morally Irrelevant Costs and Benefits

When considering issues of fairness (such as whom to save from death when only some imperiled people can be rescued), and in calculating the net benefit of infringing behavior (as threshold deontology requires), deontology typically diverges from consequentialism (including welfare economics) in yet another respect. Consequentialism usually aggregates all benefits and costs indiscriminately, whereas deontology often excludes various costs and benefits from the moral calculus (Brennan 1995; Zamir and Medina 2010: 86–93). For example, a deontologist might plausibly argue that when choosing between a course of action in which two people are saved from death and a third person is relieved from a mild headache, and a course of action where two other people are saved from death and a car is salvaged, the prospects of relieving the headache or of salvaging the car are simply irrelevant. Similarly, a deontologist may distinguish between harming someone to prevent considerable harm befalling other people, and harming someone in order to further improve other people's well-being. Contrary to standard CBA, where forgone benefits are merely a type of costs, and forgone costs are simply benefits, the deontologist might wish to exclude—or drastically discount—the enhancement of well-being as justifying infringements of constraints (Zamir 2015: 189–90).

The upshot of all this is that, as a consequentialist normative theory, welfare economics may be faulted on three major counts: the absence of constraints on pursuing the overall good; the absence of options *not* to pursue the overall good; and the consideration of morally irrelevant factors when constraint- (and option-) infringements are at stake. In other words, the first critique claims that welfare economics *allows* too much, and the second that it *demand*s too much. It fails to recognize the moral rights of people over their body, labor, and talents (cf. Alexander 2004). It legitimizes and even requires harming other people, lying, and breaking promises, to achieve desirable results (Williams 1973: 93–107; Nagel 1986: 175–88). Normative economics arguably demands that we harvest the organs of one person to save the lives of two others; that we torture the baby of a terrorist to force him to reveal information that may save lives; and so forth. Concomitantly, it requires the well-off to devote almost all of their money, and a large portion of their time and energy, to promoting the well-being of disadvantaged people around the world. It also requires self-sacrifice when the expected benefit to another person (who may be just as well-off) is only slightly more than the cost to oneself.

The concern that consequentialism justifies terrible deeds is exacerbated when the theory of human welfare underpinning the normative theory is the satisfaction of actual preferences, and even more so if preferences are measured by people's willingness to pay for their satisfaction—as welfare economics ordinarily posits. People sometimes have prejudiced, xenophobic, or even sadistic preferences, and their willingness to pay to satisfy these preferences directly depends on their affluence. Taken to their logical conclusion, these features can, in theory at least, lead to justifying “efficient” rapes, murders, and even genocide (Williams 1973: 105; Dolinko 2002).

24.3.5 Criticism of Deontology

Of course, deontological morality is also susceptible to criticism on various counts. One critique often leveled against deontological constraints—both moderate and absolutist—is that they are irrational. Assuming that it is intrinsically bad to harm someone else or to use her as a means to an end, it seems irrational to oppose such harm or use when the consequence of not doing so is a greater amount of equally severe harm or use of other people. In response to this “paradox of deontology” critique, one might argue that constraints are necessary to address people's biases, shortsightedness, and cognitive limitations (Kagan 1989: 32–9). However, this response is unpersuasive when it is certain that non-infringement of the constraint would actually increase its overall violations.

Alternatively, one might reject the implicit consequentialist premise that the desirability of actions is ultimately determined by their outcomes—that is, to divorce the right from the good. Thus, a deontologist might claim that duties arising from special relationships (such as between parents and their children), or the prohibition of certain relationships (such as that of torturer and victim) take precedence over the duty to maximize the good (Ross 1930: 16–47; McNaughton and Rawling 1998). While consequentialism judges acts by their outcome, deontology judges acts based on whether or not the actors conduct themselves in ways that maintain their moral integrity (Darwall 2003b).

Another criticism of deontology (moderate or absolutist) is the great difficulty in systematically and generally characterizing what the constraint of not-harming people actually means. Shelly Kagan (1989: 83–182), among others, has powerfully demonstrated that none of the conventional deontological distinctions—between doing harm and merely allowing it, between intending harm and merely foreseeing it, or between interference and non-interference with others' welfare—offer a coherent and intuitively acceptable criterion for establishing the constraint in all circumstances. Although some philosophers believe it is possible, nonetheless, to come up with such a systematic criterion (Kamm 1989), others concede that it may not exist, but argue that this in itself need not be a fatal flaw of deontology. It may be, instead, that different factors play different roles in different contexts (Griffin 1990; Freeman 1994).

The final critique of deontology to be mentioned here is that any non-consequentialist concern is incompatible with the (weak) Pareto principle, and therefore unacceptable. While a consequentialist normative theory that determines the goodness of outcomes

based solely on their impact on people's well-being will always favor a rule (or anything, for that matter) that makes everyone better off, a deontological theory may occasionally favor a rule that makes everyone worse off (Kaplow and Shavell 2002: 52–8). In response to this claim, deontological theories do not *necessarily* conflict with the Pareto principle. A deontological theory may plausibly qualify any constraint or option such that it does not apply whenever everyone is better off without it (Chang 2000; Craswell 2003). More fundamentally, the conflict between deontology and the Pareto principle is tautological: it merely restates the conflict between deontology and consequentialism, and as such may be regarded as highlighting the weakness of the Pareto principle (Sen 1970; White 2009, 2011: 139–43). Finally, in the context of deontological constraints, Kaplow and Shavell's argument is valid only behind an imaginary contractarian veil of ignorance: in real life, it is the *removal* of a constraint against harming people that necessarily violates the Pareto principle. Removing the constraint renders the harming of people obligatory whenever such harming produces good outcomes overall (at the expense of those being harmed).

In addition to these general criticisms of deontological morality, moderate deontology is also criticized for being incoherent, disrespectful of human dignity, arbitrary, and giving rise to strange conundrums. For example, setting thresholds is arguably incoherent: either deontological constraints have primacy over the good, in which case they should never be violated, or they can be overridden by the goodness of results, in which case one becomes a consequentialist (Davis 1991). The response to this critique is that it is perfectly coherent to maintain that (contrary to consequentialism) the goodness of outcomes is not the only factor, and that (contrary to absolutist deontology) constraints may be outweighed by enough good outcomes. Recognizing that there is more than one morally relevant factor inevitably implies that under different circumstances some factors outweigh others, and that all factors should be taken into account.

Relatedly, some argue that the very formulation of rules that determine when it is permissible to kill or torture people is disrespectful of human dignity, and thus incompatible with Kantian morality. Extreme emergencies may indeed compel one to do horrible things to prevent catastrophic outcomes. However, respect for people requires that such acts be “unprincipled, context-generated,” and performed strictly as acts of necessity (Harel and Sharon 2011). In response, it may be argued that many infringements of deontological constraints—including lying, promise breaking, or causing mild physical pain—are deemed permissible under ordinary circumstances, not merely under extreme ones. Moreover, even if one were to restrict the discussion to extreme measures such as killing and torture, if no moral principles govern the behavior of agents in extreme cases, how can agents decide whether the circumstances in question are truly extreme? How can one judge, in retrospect, whether a particular infringement was morally justified? There may be institutional and expressive reasons for refraining from explicitly authorizing the killing or torturing of innocent people by the legal system; and there may be instrumental advantages to viewing some acts as morally taboo. However, once it is accepted that constraints may justifiably be infringed under certain circumstances, there is no avoiding the need to delineate those circumstances.

Moderate deontology is also criticized for being arbitrary. It is impossible, it is claimed, to set the deontological threshold at any particular point in a non-arbitrary manner (Ellis 1992; Alexander 2000; Lowry and Peterson 2012). This is arguably due to the incomparability of consequentialism and deontology: the former is goal-oriented, committed to maximizing good outcomes, while the latter holds that some acts are intrinsically wrong. Similar allegations are made with regard to thresholds for deontological options (Alexander 1996). In response, one might point out that, although any cut-off point is indeed “arbitrary” in the sense that the threshold could have been fixed higher or lower, this is true of any pluralist normative theory involving more than one morally relevant factor (Moore 1997: 723–4; Kagan 1998: 80–1).

Finally, the discontinuity brought about by setting thresholds arguably leads to strange moral conundrums (Alexander 2000, 2011). Suppose, for example, that it is morally correct to torture a terrorist’s mother (to put pressure on the terrorist to reveal where he planted a bomb) to save the lives of n people—but not to save the lives of $n-1$ people. Must the torture be halted the moment the number of threatened people drops from n to $n-1$? And if the number of people at risk is just one short of the threshold, could the police justifiably lure one more person into the danger zone in order to legitimize the torture and save the lives of all of those at risk? At least in the legal context, where a myriad of normative, pragmatic, and institutional considerations either justify somewhat troubling, discontinuous outcomes, or lead to the establishment of more or less vague standards to be applied on a case-by-case basis, such (somewhat gimmicky) examples do not seem dispositive. The problem of transforming gradated, quantitative differences into discontinuous, qualitative outcomes would plague the legal system even if it were purely consequentialist.

24.3.6 Moderate Deontology and the Law

Although we recognize that there are challenges with deontology in general and with moderate deontology in particular, we nonetheless believe that moderate deontology is more attractive than consequentialism in general, and standard welfare economics in particular. Furthermore, for principled and pragmatic reasons, we believe that moderate deontology is particularly apposite as a decision procedure for legal policy- and decision-making. This is because moderate deontology is far more compatible with prevailing moral convictions (or *commonsense morality*) than either consequentialism or absolutist deontology. This compatibility has long been recognized by ethicists (Scheffler 1988; Kagan 1989: 1–5; Brennan 1995), and in recent years has been confirmed in numerous empirical studies of people’s moral judgments. Those studies have found that, although a small minority of people reason as consequentialists, and a large minority of people, at least initially, reason as absolutist deontologists, most people tend to be moderate deontologists (Zamir 2015: 182–8). Inasmuch as democratic theory maintains that the state should put prevailing normative judgments into effect (Balkin 2011: 64–9), the empirical findings support the adoption of moderate deontology as a framework for legal

policymaking. Furthermore, insofar as people's compliance with legal norms is enhanced if they find them morally appealing, moderate deontology should be adopted for instrumental reasons, as well (cf. Robinson and Darley 1997).²

The next section critically analyzes attempts to defend consequentialism against the critiques that it both permits and demands too much, and several proposals to modify standard economic analysis to overcome these flaws.

24.4 INDIRECT WAYS OF ADDRESSING DEONTOLOGICAL CONCERNS

24.4.1 General

Proponents of moral theories usually seek to demonstrate that the theory they espouse is in keeping with prevailing moral convictions and does not lead to morally repugnant conclusions. Notwithstanding their predilection for counterintuitive claims (Dorff and Ferzan 2009), this is basically true of consequentialists, including lawyer-economists. Several responses have been offered to the deontological critique of consequentialism, and of economic analysis in particular.³

One response is to demonstrate that the counterintuitive conclusions attributed to consequentialism are based on a flawed analysis, which disregards or underestimates relevant outcomes. A more comprehensive analysis—so the argument goes—would reveal that seemingly efficient acts that violate deontological constraints are not, in fact, efficient, and therefore unjustified on purely consequentialist grounds. Thus, for example, it has been argued that killing one person in order to use her organs to save the lives of two other individuals only *appears* to be desirable, because once one takes into account the long-term and indirect effects of such action, its costs prove to outweigh its benefits (Singer 1977, responding to Harris 1975). Similarly, the claim that market discrimination should not be counteracted because it must be efficient (because members of protected groups could have bargained to avoid inefficient discrimination) ignores the crucial problem of externalities (Donohue 1992; Iacobucci 1998). Those who suffer from discrimination cannot be expected to purchase the bigots' freedom to discriminate because the benefits of such a bargain have the hallmarks of a public good. Similarly, employment discrimination undermines the *ex ante* incentives for investment in human capital by would-be workers who assume that they would be treated as average members of their

² Due to space constraints, we do not discuss the claim that, even if moderate deontology is an appropriate private morality, it is inappropriate in the public sphere. See Goodin (1995), Sunstein and Vermeule (2005), Zamir and Medina (2010: 57–78), Wonnell (2011), and Hosein (2014).

³ The strategies briefly described below address the claim that consequentialism lacks constraints. Similar strategies are used to answer the critique that it lacks options, which we will not discuss here due to space constraints. See Zamir and Medina (2010: 33–40).

group, rather than according to their specific qualities. However, such analyses are often strained (Williams 1973: 100); they still leave *some* instances where consequentialism yields abhorrent conclusions, and from a deontological perspective, even when they are successful, they avoid abhorrent conclusions for the wrong reasons.

A more promising strategy is to move from *act-consequentialism* to *rule-consequentialism*. According to the latter, even if killing one person to save the lives of two others may bring about good outcomes overall, given the fallibilities of human decision-making it may still be advisable to adopt an absolute, or almost absolute, prohibition of killing people. Such a rule may bring about better results overall, even if in some particular cases it may yield sub-optimal outcomes. To take another example, one may argue that since markets are rarely fully competitive, anti-discrimination legislation is essential to producing the outcome that would result in a fully competitive market (Donohue 1986). Without delving into the debate over rule-consequentialism (Smart 1967; Hooker 2000; Parfit 2011), such an argument does not redeem standard economic analysis, which is distinctively act-consequentialist. At most, it provides a consequentialist justification for the adoption—by legal policymakers and others—of a decision strategy that incorporates deontological constraints and options (Zamir and Medina 2010: 24–7).

Another possibility is to adopt a sophisticated theory of good outcomes that takes into account considerations of rights, agent-relativity, how outcomes are brought about, and so forth—what David McNaughton and Piers Rawling (1991) have vividly described as the *consequentialist vacuum cleaner* (see also Sen 1982, 1985; Portmore 2009). However, this strategy is open to question (Schroeder 2007), and in any event is irrelevant to economic analysis, which is based on a simple welfarist theory of the good—namely, preference satisfaction. The closest one can get to defending welfare economics by improving its underlying theory of the good (without dramatically changing its basic features) is by shifting from the satisfaction of *actual* preferences—the traditional measure of well-being in economic analysis—to satisfying *ideal* ones. Welfare economics can, in principle, disregard or discount misinformed, biased, prejudiced, and irrational preferences (Adler and Posner 2006: 124–53; Lowry and Peterson 2012). While such a strategy may mitigate some of the abhorrent implications of economic analysis (such as legitimizing “efficient” rapes), it still does not answer many other deontological concerns, such as the immorality of killing one innocent person to save the lives of two.

At this point, one might give up on the attempt to address deontological concerns within economic analysis.⁴ One might, for example, propose that academic jurists and legal policymakers consider non-efficiency arguments separately from economic analysis (Adler and Posner 2006: 154–8). Some commentators, however, refuse to give up so easily, proposing other strategies for improving economic analysis. The more popular strategy among lawyer-economists is to treat moral judgments as preferences whose fulfillment should be maximized along with any other.

⁴ Even more drastically, one might put aside the deontology-consequentialism debate altogether and adopt a virtue ethics perspective on economic analysis instead (Baker and White 2016). This proposal lies beyond the scope of our discussion.

24.4.2 Moral Preferences

One common response to the anti-consequentialist critique of welfare economics is to incorporate people's moral attitudes—including the preference for “fairness” and prohibitions against harming others—into CBA (Zerbe 2001; Kaplow and Shavell 2002: 431–6; Calabresi 2016; Posner and Sunstein 2017). The various proposals offered along these lines share two common convictions. On the one hand, they accept that moral considerations should not be excluded from normative economic analysis; on the other hand, they interpret the fundamental ideal of economic analysis, namely the commitment to normative neutrality, as prohibiting the incorporation of morality into the analysis (Ulen 2015). The alleged solution is to draw on the reflection of moral judgments in the empirical world. This approach raises two main difficulties. Methodologically, it is difficult to identify and aggregate “preferences” about moral issues. Substantively, it is doubtful that normative questions can be resolved by aggregating people's preferences. In what follows we critically evaluate the main proposals made along these lines.

With regard to the methodological issue, one approach is to suggest that moral preferences can and should be inferred from institutional choices. Thus, Guido Calabresi (2016) argues that economic insights can not only enrich legal analysis, but that close examination of legal norms and institutions can enrich economic analysis as well. When existing legal norms and institutions do not conform to the predictions of conventional economic analysis, this may be indicative that conventional analysis is lacking. The missing element, so argues Calabresi, is often the recognition of the existence of certain widespread moral attitudes. Therefore, in order to live up to their proclaimed neutrality with regard to the objects and content of people's preferences, economists should incorporate those attitudes into their analyses. Because these attitudes are usually reflected in the legal system, the law is an invaluable source for identifying them.

This approach aims to resolve the practical difficulties of aggregating moral preferences by relying on the outcomes of the legal and political processes, assuming that these social institutions produce the required aggregation. However, this approach raises considerable doubts. The idea that the current legal system adequately reflects the prevailing tastes and values of the populace is questionable. If it is based on the premise that legal policymakers—legislators, public officials, and judges—strive to fulfill the tastes and values of the general public, this premise is very problematic. Even if one relaxes the simplistic assumptions underpinning public choice theory (which are especially moot in the context of judicial decision-making) and maintains that public officials are motivated in part by a desire to promote the general good and by other non-selfish commitments, it is still a far cry from assuming that the political mechanism adequately aggregates the populace' moral judgments. Quite often, it does not. Moreover, while politicians and other public officials sometimes attempt to identify and quantify the prevailing tastes and values in society, most legislative, administrative, and certainly judicial decisions are not made on this basis. Even when legal policymakers know, or believe they know, what the most popular decision would be, they often opt for less popular ones.

Alternatively, one may assume that the legal system reflects the prevailing tastes and values not (or not only) because public officials are able and eager to promote the aggregated tastes and values of the population as a whole, but because their own values are similar to those of the population as a whole; otherwise, presumably, they would not have been elected for, or appointed to, their public positions. However, even if such reflection were theoretically possible, it appears to be a rather romantic view of the political process—even in liberal democracies. Ordinarily, a very large percentage of eligible voters do not vote, and those who do often make their choices based on the candidates' positions on a limited range of issues—yet, once elected, these officials decide a much broader range of issues. Finally, if in fact the political process is an adequate mechanism for aggregating preferences and normative judgments, than—almost by definition—*any* existing legal regime in any society may be said to maximize the aggregate social welfare, and is therefore desirable (Malcai 2017; Zamir 2017).

Another, more popular approach proposes that we identify and aggregate moral preferences by measuring people's willingness to pay for their satisfaction (Zerbe 2001; Posner and Sunstein 2017). To measure the "existence value" or "non-use" value of things such as environmental protection and wildlife conservation, economists have developed various techniques, including *contingent valuation methods* (CV or CVM), based on public polls (Mitchell and Carson 1989; Champ et al. 2017; Carson 2011). For example, John Donohue (1987) argued that for anti-discrimination laws to be efficient, it is sufficient that every American is willing to pay as little as five dollars annually to live in a society that curbs racial discrimination. Such preferences cannot be satisfied through voluntary bargaining, because the benefits of such a bargain are a public good, and can thus justify the government's intervention.

In principle, such methods might be used to elicit information about people's attitudes to human rights and morality in general. However, such methods have also been the subject of much criticism. First, it is possible that preferences regarding non-market goods, especially vague or distant ones, are simply not sufficiently consistent or clear-cut to be measurable (Schkade and Payne 1993). Even when preferences do exist, CV studies are often inaccurate and inconsistent. Variations in the order of goods or options being presented result in different valuations—as do variations in the framing of the proposed change, and whether or not the goods are perceived as replaceable. Moreover, when people state how much they are willing to pay for a certain public or merit good, they tend to ignore numerous other worthwhile goals that they might wish to advance (and their limited budget). People's answers also depend on how much they enjoy the feeling of (hypothetical) giving, or alternatively how much they would like to gain a free ride on other people's contributions, which adds "noise" to the results (Diamond and Hausman 1993; Carson 2012).

Beyond such methodological difficulties of gauging people's moral judgments, aggregating moral "tastes and values" raises principled objections. The fulfillment of preferences or desires may be regarded as a measure of people's welfare or utility. As long as one strives to maximize human welfare, maximizing the fulfillment of people's preferences (such as a preference for vanilla ice cream) appears to be sensible. However, when one moves from human welfare to normative questions, preferences are not an

acceptable criterion.⁵ Judging the morality of a given act or rule based on all relevant factors (including, for example, whether it involves active or intended harm to others, or disparate treatment of similarly situated people) must not be confused with one such factor, namely the act's or rule's effect on the well-being of the person making the judgment (Adler and Posner 2000). There is a fundamental difference between preferences and normative judgments (Malcai 2017). Whether a judgment is right or not depends on its justifications, not on the number of its supporters or the intensity of their support (Kornhauser 2003; Adler and Posner 2000; Herzog 2000). Thus, while there is a good utilitarian reason to sum up people's preferences in the narrow sense, it is normatively unsound to sum up people's moral judgments or values.

Moreover, indiscriminate incorporation of people's moral judgments in CBA may not only fail to bridge the gap between welfare economics and moderate deontology—it may actually widen it, if enough people have sufficiently intense anti-social “preferences.” For example, if the discriminatory preferences of employers, coworkers, and customers are included in the social welfare calculation and aggregated based on people's willingness to pay for their satisfaction, it may be difficult to justify anti-discrimination laws (Epstein 1992: 486–8). While ignoring such preferences as being irrational or socially unacceptable may well be warranted (Landes 1968; Stein 2003), doing so would mean that the analysis is no longer value-free. Finally, if one treats normative judgments as preferences to be maximized, and if enough people object to consequentialism, welfarism, utilitarianism, or measuring people's welfare by the extent to which their preferences are fulfilled, then aggregating these “preferences” may arguably lead to abandoning normative economics altogether.⁶ Confounding preferences and value judgments within a normative economic analysis is therefore methodologically and conceptually problematic, and potentially self-defeating.

24.5 INTEGRATING THRESHOLD CONSTRAINTS AND OPTIONS WITH ECONOMIC ANALYSIS OF LAW

Given the methodological and substantive difficulties of treating deontological constraints and options as preferences, we have proposed an analytical framework that does away with the purely consequentialist underpinnings of welfare economics (Zamir and Medina 2010; see also Hosein 2016). We suggested that the permissibility of any given act or rule that infringes a deontological constraint (or option) be determined by means

⁵ It does not follow, however, that choices and decisions that are motivated by moral commitments should not be taken into account when explaining or predicting people's behavior (Hausman 2005; see also Sen 1977, Brennan and Eusepi 2009, and Zamir 2017). It is one thing to explain and predict human behavior, and another thing entirely to formulate legal policies.

⁶ In fact, as previously noted, most people reason as moderate deontologists rather than consequentialists, and there appears to be widespread aversion to CBA (Viscusi 2000).

of *threshold functions*. The proposed framework calls on legal policymakers to tackle the pertinent normative issues head-on as normative issues. Accordingly, compliance with deontological constraints is not something to be maximized along with other objects of people's preferences. Rather, the prohibition on active or intended harming of innocent people, to take one example, is a constraint on maximizing aggregate utility.

Threshold functions sometimes involve the thorny task of monetizing deontological constraints and options—but on other occasions, they do not. For example, if under a given threshold function the only benefit that may ever justify the active/intentional killing of x innocent people is the saving of the lives of at least m times x people (or m times x , plus n people), then no monetization of human life or of m or n is necessary. Even without inserting actual numbers into threshold functions, we believe that, much like many other economic models constructed by lawyer-economists, they can clarify the pertinent issues and their interrelationships. When threshold functions do require monetization of constraints and options to incorporate them into deontologically constrained CBA, they face the objection of inappropriate *monetization* of incommensurable objects, which is arguably impossible or incompatible with the spirit of deontology (Kelman 1981; Wonnell 2011; Lowry and Peterson 2012). But they do not face the objection to *commodification* of non-market objects, as the proposal does not treat moral constraints and options as goods that may be traded in the market (Zamir and Medina 2010: 108–16; Zamir 2017).

As previously noted, when determining the permissibility of actions and rules that infringe a deontological constraint, deontological morality usually excludes all sorts of costs and benefits from the calculation. Contrary to standard CBA, when deontological morality considers, for example, the permissibility of killing some people to save others (or as a side-effect of saving the latter), it may (or may not, depending on the theory's details) deem body injuries, damage to property, and pecuniary gains and losses to be irrelevant. Therefore, in constructing threshold functions, one should also determine the type of costs and benefits that should be taken into account. Among the possible excluders of costs and benefits are those derived from the notion of “lexical priority,” a moral judgment regarding the treatment of chronologically remote or probabilistic costs and benefits (and other factors) (Zamir and Medina 2010: 86–93).

Consider, once again, the issue of marketplace discrimination. Deontological morality plausibly maintains that discrimination is inherently wrong, as it violates the prohibition against harming the people who are discriminated against (Dworkin 2000; Hellman 2008; Moreau 2010; Roberts 2016). Assessing any anti-discrimination legislation requires, first, to prioritize between two conflicting constraints. On the one hand, there is the constraint against harming people by discriminating against them, thereby diminishing their autonomy, offending their dignity, and depriving them of various material and non-material goods. On the other hand, there is the constraint against limiting people's freedom, including the freedom to choose with whom to interact and contract. When the latter constraint trumps the former—such as in the sphere of intimate relationships—proscribing discrimination is unjustified. When the constraint against discrimination takes priority—such as, plausibly, in the labor market—a further question

arises, namely whether the costs involved in securing equal treatment are high enough to override the constraint. Deontologically constrained CBA may contribute to understanding dilemmas of this sort. Moreover, it would not only be better at capturing the relevant normative questions, but would also be in line with current legal doctrines.

The moral constraint against discrimination means that prohibiting discrimination on grounds such as race and sex can be justified even if its social cost exceeds its benefit. This is not to say that welfare analysis is irrelevant when evaluating the desirability of anti-discrimination legislation or implementing concrete legal provisions. A sufficiently high net cost may well justify infringing the moral constraint against discrimination, and a corresponding limitation of the scope of anti-discrimination legal norm (or an exception to it). To assess the permissibility of a given type of discrimination, one needs a threshold function that determines the size of the threshold and the pertinent types of costs and benefits. The size of the threshold may depend on the form of discrimination, which serves as a proxy for the extent of the harm. Plausibly, the moral constraint against the use of an employment practice that causes a *disparate impact* on the basis of race, color, religion, sex, or national origin is significantly less strict than the constraint against refusing to hire or otherwise discriminating against an individual based directly on one of these grounds.⁷ Such differences may justify variations in the applicable threshold. As for the types of costs and benefits that are relevant to determining whether or not a threshold is met, one key question is whether discriminatory preferences and prejudices of the employer or of the firm's customers or other employees should be taken into account. As we have seen, disregard for antisocial preferences is possible even within a purely consequentialist theory (and has in fact been advocated by some economists). However, it is more likely to be part of a deontologically-constrained CBA. The anti-discrimination constraint implies that satisfying racial or sexist preferences should not constitute part of the social good (Zamir and Medina 2010: 240–55).

Threshold functions do not provide instant or easy solutions to legal problems. They may, however, point legal policymakers to the crucial issues at stake, and combine the analytical rigor of economic analysis with moderate deontological morality. Constructing a threshold function, deciding about the size and shape of the threshold, and determining which cost and benefits should be excluded from the analysis, all require judgment, but—contrary to what some critics say—they are not arbitrary (Zamir and Medina 2010: 53; Hosein 2016). In addition to highlighting the relevant normative issues and tackling them as such, this proposal makes explicit trade-offs between values that, at first blush, appear to be incommensurable. While open public discourse about such trade-offs may be undesirable for various reasons, such trade-offs are often inevitable, and conducting them in a rigorous manner may help legal policymakers when making their decisions.

Just as non-consequentialist normative judgments may be incorporated into economic analysis, it may be possible, in principle, to reframe CBA so that it incorporates non-

⁷ According to a common legal doctrine, an apparently neutral provision, criterion or practice may nevertheless be considered as discriminatory, if it has a distinctive adverse effect—or disparate impact—on members of a protected group.

welfarist and non-utilitarian concerns—although considerably more work must be done to formulate a workable framework of this sort. The crucial point is that such analytical frameworks would not treat moral issues as mere elements of human welfare to be maximized, but rather as concerns that are important in their own right. The scope, content, and weight of these concerns are debatable—and should indeed be debated—but true normative and legal deliberation cannot circumvent them.

24.6 CONCLUSION

The more jurists and legal policymakers employ economic analysis in virtually all legal spheres, and the more public authorities use CBA as a decision tool, the more it becomes apparent that disregard of standard economic analysis for non-consequentialist moral concerns is profoundly troubling, for principled as well as for practical reasons. Many believe that it is impossible to integrate non-consequentialist values and economic analysis within a single analytical framework. Others, however, argue that, while the challenges facing such an integration are significant, they are not insurmountable. This chapter offered a non-exhaustive critical survey of the main attempts to overcome these challenges.⁸ Notwithstanding their limitations and difficulties, we believe that such attempts are preferable to a situation where ethicists, including deontologically inclined jurists, and economists, including lawyer-economists, keep talking past each other.

ACKNOWLEDGMENTS

We thank Tal Mendelson, Asor Watzman, and Roi Yair for their excellent research assistance.

REFERENCES

- Adler, Matthew D., and Eric A. Posner. 2000. "Implementing Cost-Benefit Analysis When Preferences Are Distorted." *Journal of Legal Studies* 29: 1105–47.
- Adler, Matthew D., and Eric A. Posner. 2006. *New Foundations of Cost-Benefit Analysis*. Cambridge, MA: Harvard University Press.
- Alexander, Larry. 1996. "Affirmative Duties and the Limits of Self-Sacrifice." *Law and Philosophy* 15: 65–74.
- Alexander, Larry. 2000. "Deontology at the Threshold." *San Diego Law Review* 37: 893–912.
- Alexander, Larry. 2004. "The Jurisdiction of Justice: Two Conceptions of Political Morality." *San Diego Law Review* 41: 949–66.
- Alexander Larry. 2011. "Deontological Constraints in a Consequentialist World: A Comment on Law, Economics, and Morality." *Jerusalem Review of Legal Studies* 3: 75–90.

⁸ See also Sunstein (2014) and Peterson (2017).

- Alexander Larry. 2016. "The Means Principle." In Kimberly Kessler Ferzan and Stephen J. Morse (eds.), *Legal, Moral, and Metaphysical Truths: The Philosophy of Michael S. Moore* (Oxford: Oxford University Press), pp. 251–64.
- Baker, Jennifer A., and Mark D. White (eds.). 2016. *Economics and the Virtues: Building a New Moral Foundation*. Oxford: Oxford University Press.
- Balkin, Jack M. 2011. *Living Originalism*. Cambridge, MA: Harvard University Press.
- Bennett, Jonathan. 1995. *The Act Itself*. Oxford: Oxford University Press.
- Brennan, Geoffrey, and Giuseppe Eusepi. 2009. "Value and Values, Preferences and Price: An Economic Perspective on Ethical Questions." In Geoffrey Brennan and Giuseppe Eusepi (eds.), *The Economics of Ethics and the Ethics of Economics: Values, Markets and the State* (Cheltenham, UK: Edward Elgar), pp. 14–31.
- Brennan, Samantha. 1995. "Thresholds for Rights." *Southern Journal of Philosophy* 33: 143–68.
- Calabresi, Guido. 1991. "The Pointlessness of Pareto: Carrying Coase Further." *Yale Law Journal* 100: 1211–37.
- Calabresi, Guido. 2016. *The Future of Law and Economics: Essays in Reform and Recollection*. New Haven, CT: Yale University Press.
- Carson, Richard T. 2011. *Contingent Valuation: A Comprehensive Bibliography and History*. Cheltenham, UK: Edward Edgar Publishing.
- Carson, Richard T. 2012. "Contingent Valuation: A Practical Alternative when Prices Aren't Available." *Journal of Economic Perspectives* 26: 27–42.
- Champ, Patricia A., Kevin J. Boyle, and Thomas C. Brown (eds.). 2017. *A Primer on Nonmarket Valuation*. 2nd edn. Dordrecht: Springer.
- Chang, Howard F. 2000. "A Liberal Theory of Social Welfare: Fairness, Utility, and the Pareto Principle." *Yale Law Journal* 110: 173–235.
- Chang, Ruth (ed.). 1997. *Incommensurability, Incomparability, and Practical Reason*. Cambridge, MA: Harvard University Press.
- Cooter, Robert, and Thomas Ulen. 2012. *Law and Economics*. 6th edn. Boston, MA: Prentice Hall.
- Craswell, Richard. 2003. "Kaplow and Shavell on the Substance of Fairness." *Journal of Legal Studies* 32: 245–75.
- Darwall, Stephen. 2003a. "Introduction." In Stephen Darwall (ed.), *Deontology* (Malden, MA: Blackwell Publishing, pp. 1–7.
- Darwall, Stephen. 2003b. "Agent-Centered Restrictions from the Inside Out." In Stephen Darwall (ed.), *Deontology* (Malden, MA: Blackwell Publishing, pp. 112–38.
- Davis, Nancy (Ann). 1991. "Contemporary Deontology." In Peter Singer (ed.), *A Companion to Ethics* (Oxford: Blackwell Publishers), pp. 205–18.
- Degrazia, David (ed.). 1996. *Taking Animals Seriously: Mental Life and Moral Status*. Cambridge: Cambridge University Press.
- Diamond, Peter A., and Jerry A. Hausman. 1993. "On Contingent Valuation Measurement of Nonuse Values." In Jerry A. Hausman (ed.), *Contingent Valuation: A Critical Assessment* (Amsterdam: Elsevier Science), pp. 3–38.
- Dolinko, David. 2002. "The Perils of Welfare Economics: Book Review of *Fairness versus Welfare*, by Louis Kaplow & Steven Shavell." *Northwestern University Law Review* 97: 351–93.
- Donohue, John J., III. 1986. "Is Title VII Efficient?" *University of Pennsylvania Law Review* 134: 1411–31.
- Donohue III, John J. 1987. "Further Thoughts on Employment Discrimination Legislation: A Reply to Judge Posner." *University of Pennsylvania Law Review* 136: 523–51.

- Donohue, John J. III. 1992. "Advocacy Versus Analysis in Assessing Employment Discrimination." *Stanford Law Review* 44: 1583–614.
- Dorff, Michael B. 2002. "Why Welfare Depends on Fairness: A Reply to Kaplow and Shavell." *Southern California Law Review* 75: 847–99.
- Dorff, Michael B., and Kimberly Kessler Ferzan. 2009. "Is there a Method in the Madness? Why Creative and Counterintuitive Proposals Are Counterproductive." In Mark D. White (ed.), *Theoretical Foundations of Law and Economics* (Cambridge: Cambridge University Press, pp. 21–40.
- Dworkin, Ronald. 2000. *Sovereign Virtue: The Theory and Practice of Equality*. Cambridge, MA: Harvard University Press.
- Ellis, Anthony. 1992. "Deontology, Incommensurability and the Arbitrary." *Philosophy and Phenomenological Research* 52: 855–75.
- Epstein, Richard A. 1992. *Forbidden Grounds: The Case Against Employment Discrimination Laws*. Cambridge, MA: Harvard University Press.
- Foot, Philippa. 1978. "The Problem of Abortion and the Doctrine of the Double Effect." In *Virtues and Vices and Other Essays in Moral Philosophy* (Berkeley, CA: University of California Press), pp. 19–32.
- Freeman, Samuel. 1994. "Utilitarianism, Deontology, and the Priority of Right." *Philosophy and Public Affairs* 23: 313–49.
- Goodin, Robert E. 1995. *Utilitarianism as Public Philosophy*. Cambridge: Cambridge University Press.
- Griffin, James. 1986. *Well-Being: Its Meaning, Measurement, and Moral Importance*. New York: Oxford University Press.
- Griffin, James. 1990. "The Limits of Morality by Shelly Kagan." *Mind* 99: 128–31.
- Harel, Alon, and Assaf Sharon. 2011. "'Necessity Knows No Law': On Extreme Cases and Uncodifiable Necessities." *University of Toronto Law Journal* 61: 845–65.
- Harris, John. 1975. "The Survival Lottery." *Philosophy* 50: 81–7.
- Hausman, Daniel M. 2005. Sympathy, Commitment, and Preference. *Economics and Philosophy* 21: 33–50.
- Hausman, Daniel, Michael McPherson, and Debra Satz. 2017. *Economic Analysis, Moral Philosophy, and Public Policy*. 3rd edn. New York: Cambridge University Press.
- Hellman, Deborah. 2008. *When Is Discrimination Wrong?* Cambridge, MA: Harvard University Press.
- Herzog, Don. 2000. "Externalities and Other Parasites." *University of Chicago Law Review* 67: 895–923.
- Hooker, Brad. 2000. *Ideal Code, Real World*. Oxford: Oxford University Press.
- Hosein, Adam O. 2014. "Doing, Allowing, and the State." *Law and Philosophy* 33: 235–64.
- Hosein, Adam O. 2016. "Deontology and Natural Hazards." In Paolo Gardoni, Colleen Murphy, and Arden Rowell (eds.), *Risk Analysis of Natural Hazards: Interdisciplinary Challenges and Integrated Solutions* (Cham, Switzerland: Springer International Publishing), pp. 137–53.
- Howard-Snyder, Frances. 2016. "Doing vs. Allowing Harm." In Edward N. Zalta (ed.), *Stanford Encyclopedia of Philosophy*, at <https://plato.stanford.edu/archives/win2016/entries/doing-allowing/>.
- Iacobucci, Edward M. 1998. "Antidiscrimination and Affirmative Action Policies: Economic Efficiency and the Constitution." *Osgood Hall Law Journal* 36: 293–337.
- Kagan, Shelly. 1989. *The Limits of Morality*. Oxford: Oxford University Press.
- Kagan, Shelly. 1998. *Normative Ethics*. Boulder, CO: Westview Press.

- Kamm, Frances M. 1989. "Harming Some to Save Others." *Philosophical Studies* 57: 227–60.
- Kamm, Frances M. 1993. *Morality, Mortality, Vol. I: Death and Whom to Save From It*. New York: Oxford University Press.
- Kamm, Frances M. 2016. *The Trolley Problem Mysteries*. Eric Rakowski (ed.). New York: Oxford University Press.
- Kaplow, Louis, and Steven Shavell. 2002. *Fairness versus Welfare*. Cambridge, MA: Harvard University Press.
- Kelman, Steven. 1981. "Cost-Benefit Analysis: An Ethical Critique." *AEI Journal on Government and Society Regulation* (Jan./Feb.): 33–40.
- Keren, Gideon, and George Wu (eds.). 2015. *The Wiley Blackwell Handbook of Judgment and Decision Making*. Chichester, UK: Wiley-Blackwell.
- Kornhauser, Lewis A. 2000. "On Justifying Cost-Benefit Analysis." *Journal of Legal Studies* 29: 1037–57.
- Kornhauser, Lewis A. 2003. "Preference, Well-Being, and Morality in Social Decisions." *Journal of Legal Studies* 32: 303–29.
- Landes, William. 1968. "The Economics of Fair Employment Laws." *Journal of Political Economy* 76: 507–52.
- Lowry, Rosemary, and Martin Peterson. 2012. *Cost-Benefit Analysis and Non-Utilitarian Ethics, Politics, Philosophy, and Economics* 11: 258–79.
- Malcai, Ofer. 2017. "The Alternative Futures of Law and Economics: Comments on Guido Calabresi, *The Future of Law & Economics: Essays in Reform and Recollection*." *Jerusalem Review of Legal Studies* 16: 83–100.
- McIntyre, Alison. 2001. "Doing Away with Double Effect." *Ethics* 111: 219–55.
- McIntyre, Alison. 2014. "Doctrine of Double Effect." In Edward N. Zalta (ed.), *Stanford Encyclopedia of Philosophy*, at <https://plato.stanford.edu/archives/win2014/entries/double-effect/>.
- McNaughton, David, and Piers Rawling. 1991. "Agent-Relativity and the Doing Happening Distinction." *Philosophical Studies* 63: 167–85.
- McNaughton, David, and Piers Rawling. 1998. "On Defending Deontology." *Ratio* 11: 37–54.
- Mikhail, John. 2011. *Elements of Moral Cognition: Rawls' Linguistic Analogy and the Cognitive Science of Moral and Legal Judgment*. New York: Cambridge University Press.
- Mitchell, Robert C., and Richard T. Carson (eds.). 1989. *Using Surveys to Value Public Goods: The Contingent Valuation Method*. Washington, DC: Resources for the Future.
- Moore, Michael S. 1997. *Placing Blame*. Oxford: Oxford University Press.
- Moreau, Sophia. 2010. "What is Discrimination?" *Philosophy and Public Affairs* 38: 143–79.
- Mulgan, Tim. 2001. *The Demands of Consequentialism*. Oxford: Oxford University Press.
- Nagel, Thomas. 1986. *The View from Nowhere*. New York: Oxford University Press.
- Parfit, Derke. 2011. *On What Matters*. Oxford: Oxford University Press.
- Peterson, Martin. 2017. *The Ethics of Technology: A Geometric Analysis of Five Moral Principles*. New York: Oxford University Press.
- Portmore, Douglas W. 2009. "Consequentializing." *Philosophy Compass* 4: 329–47.
- Posner, Eric A., and Cass R. Sunstein. 2017. "Moral Commitments in Cost-Benefit Analysis." Manuscript, available at: <https://ssrn.com/abstract=2930450>.
- Posner, Richard A. 1987. "The Efficiency and the Efficacy of Title VII." *University of Pennsylvania Law Review* 136: 513–21.
- Posner, Richard A. 2001. *Frontiers of Legal Theory*. Cambridge, MA: Harvard University Press.
- Posner, Richard A. 2014. *Economic Analysis of Law*. 9th edn. New York: Aspen Publishers.

- Posner, Richard A. 2015. "Norms and Values in the Economic Approach to Law." In Aristides N. Hatzis and Nicholas Mercuro (eds.), *Law and Economics: Philosophical Issues and Fundamental Questions* (Abingdon, UK: Routledge), pp. 1–15.
- Rawls, John. 1999. *A Theory of Justice*. Rev. ed. Cambridge, MA: Harvard University Press.
- Roberts, Jessica L. 2016. "Rethinking Employment Discrimination Harms." *Indiana Law Journal* 91: 35–454.
- Robinson, Paul H., and John M. Darley. 1997. "The Utility of Desert." *Northwestern University Law Review* 91: 453–99.
- Ross, W. David. 1930. *The Right and the Good*. Philip Straton-Lake (ed.). Oxford: Oxford University Press (2002 edition).
- Scheffler, Samuel. 1988. "Introduction." In Samuel Scheffler (ed.), *Consequentialism and Its Critics* (Oxford: Oxford University Press), pp. 1–13.
- Scheffler, Samuel. 2001. *Boundaries and Allegiances: Problems of Justice and Responsibility in Liberal Thought*. Oxford: Oxford University Press.
- Scheffler, Samuel. 2004. "Doing and Allowing." *Ethics* 114: 215–39.
- Schkade, David A., and John W. Payne. 1993. "Where do the Numbers Come From?" In Jerry A. Hausman (ed.), *Contingent Valuation: A Critical Assessment* (Amsterdam: Elsevier Science), pp. 271–303.
- Schroeder, Mark. 2007. "Teleology, Agent-Relative Value, and 'Good.'" *Ethics* 117: 265–95.
- Sen, Amartya. 1970. "The Impossibility of a Paretian Liberal." *Journal of Political Economy* 78: 152–7.
- Sen, Amartya. 1977. "Rational Fools: A Critique of the Behavioral Foundations of Economic Theory." *Philosophy and Public Affairs* 6: 317–44.
- Sen, Amartya. 1982. "Rights and Agency." *Philosophy and Public Affairs* 11: 3–39.
- Sen, Amartya. 1985. "Well-Being, Agency and Freedom: Dewey Lectures." *Journal of Philosophy* 82: 169–221.
- Singer, Peter. 1977. "Utility and the Survival Lottery." *Philosophy* 52: 218–22.
- Smart, J.J.C. 1967. "Extreme and Restricted Utilitarianism." In Philippa Foot (ed.), *Theories of Ethics* (Oxford: Oxford University Press), pp. 171–83.
- Stein, Michael A. 2003. "The Law and Economics of Disability Accommodations." *Duke Law Journal* 53: 79–191.
- Sunstein, Cass R. 2014. *Valuing Life: Humanizing the Regulatory State*. Chicago, IL: Chicago University Press.
- Sunstein, Cass R., and Adrian Vermeule. 2005. "Is Capital Punishment Morally Required? Acts, Omissions, and Life-Life Tradeoffs." *Stanford Law Review* 58: 703–50.
- Steinbock, Bonnie, and Alastair Norcross (eds.). 1994. *Killing and Letting Die*. 2nd edn. New York: Fordham University Press.
- Thomson, Judith Jarvis. 1986. "Some Ruminations on Rights." In *Rights, Restitution, and Risk: Essays in Moral Theory* (Cambridge, MA: Harvard University Press), pp. 49–65.
- Thomson, Judith Jarvis. 1985. "The Trolley Problem." *Yale Law Journal* 94: 1395–415.
- Ulen, Thomas. 2015. "Law and Economics, the Moral Limits of the Market, and Threshold Deontology." In Aristides N. Hatzis and Nicholas Mercuro (eds.), *Law and Economics: Philosophical Issues and Fundamental Questions* (Abingdon, UK: Routledge), pp. 203–25.
- Vassel, Ron A. 1994. "The Americans with Disabilities Act: The Cost, Uncertainty and Inefficiency." *Journal of Law and Commerce* 13: 397–411.
- Viscusi, W. Kip. 2000. "Corporate Risk Analysis: A Reckless Act?" *Stanford Law Review* 52: 547–97.

- White, Mark D. 2009. "Pareto, Consent, and Respect for Dignity: A Kantian Perspective." *Review of Social Economy* 67: 49–70.
- White, Mark D. 2011. *Kantian Ethics and Economics: Autonomy, Dignity, and Character*. Stanford, CA: Stanford University Press.
- Williams, Bernard. 1973. "A Critique of Utilitarianism." In J.J.C. Smart and Bernard Williams, *Utilitarianism—For and Against* (Cambridge: Cambridge University Press), pp. 75–150.
- Wonnell, Christopher T. 2011. "Deontology, Thresholds, and Efficiency." *Legal Theory* 17: 301–17.
- Zamir, Eyal. 2015. *Law, Psychology, and Morality: The Role of Loss Aversion*. New York: Oxford University Press.
- Zamir, Eyal. 2017. "Tastes, Values, and the Future of Law and Economics." *Jerusalem Review of Legal Studies* 16: 101–23.
- Zamir, Eyal, and Barak Medina. 2010. *Law, Economics, and Morality*. New York: Oxford University Press.
- Zamir, Eyal, and Doron Teichman. 2018. *Behavioral Law and Economics*. New York: Oxford University Press.
- Zerbe Jr., Richard O. 2001. *Economic Efficiency in Law and Economics*. Cheltenham, UK: Edward Elgar.

CHAPTER 25

THE ETHICS AND ECONOMICS OF ECOLOGICAL JUSTICE

DAVID SCHMIDTZ

25.1 INTRODUCTION

NOT many animals ever think about justice (and so far as we know plants never think about anything), but humans do. What is justice? The question sounds impossibly difficult, but one thing we know is that the question matters.

Strikingly, simply pointing out that the question matters also points to a partial answer. Namely, justice is a response to a problem, and the problem to which justice is a response is a problem that matters. Specifically, we can infer from observed linguistic practice that when people talk about justice, they are talking about what people in a community ought to be able to expect from each other. What should people think of as their due? What are they due from each other? We have many theories, but many of our theories are childish—ideas that siblings develop in early childhood when they think their feelings are decisively important.

What, then, is the adult problem to which justice is a response? What would it be like to have grown-up views about what people ought to be able to expect from each other? What should life in a community be like? The most adult answers presuppose a prior question, namely, what are the realistic possibilities? What can people realistically expect from each other? What could people live with, given that adults need to outgrow thinking of themselves as the center of the universe, and need to accommodate themselves to a world that for good reason does not see them that way?

If we are to emerge from the process of growing up with an adult conception of justice, it will be a conception of our place and of our due, alongside a conception of what

other people are due, within a community that has a logic of its own. That logic will be described by economics (among other social sciences) and by ecology. Realistic possibilities are implications of those logics. Our adult views about ethics and justice will be disciplined by those logics.

What would it be like to have a view of how life in a community ought to go, and of what we reasonably ought to be able to expect from each other, that is disciplined by the logic of economics and of ecology? One kind of non-question-begging answer would first acknowledge, as an economist or ecologist might, that what people expect from each other—above all whether they trust each other—affects whether their communities are places where people thrive. Then it would look for evidence regarding whether people have better prospects by virtue of having one set of expectations rather than another.

What makes some expectations more realistically conducive to human flourishing in actual historical practice than others? Ecological reasoning, like economic reasoning, is reasoning about competition, scarcity, consequences predictable and consequences unintended, plus general principles describing the logic of how systems respond to attempts to manipulate them. Here are several ideas that would make a conception more or less ecological.¹

25.2 TOWARD CONCEIVING OF JUSTICE AS REALISTICALLY ECOLOGICAL

First, *justice arises from conflict*. More specifically, our sense of justice, as a community and as individuals, emerges from conflict. Litigants go to courts of justice hoping that a judge will resolve their conflict.

Suppose an emerging constellation of mutual expectations does indeed help people to resolve or avoid conflict. Suppose their community is where they have a chance to live the life they want, or at least that it is no one's fault if they cannot. That would go a long way toward everyone involved having reason to call those expectations just.

Second, *justice manages traffic*. The glory of western civilization is not the extent to which we reached consensus but the extent to which we made consensus unnecessary. Freedom of religion may be western civilization's greatest achievement, but it had nothing to do with reaching consensus on which religion is correct. The triumph was in learning not to need that kind of consensus. We needed to reach consensus on jurisdiction, on who gets to decide. Or, to use my preferred metaphor—which is not merely a metaphor but itself a crucial challenge in all kinds of ways—we need to manage traffic, and need to do it without telling anyone that their destination is second class.

¹ For some existing literature on environmental justice (Stone 1996; Bullard 1990; Norton 1991; Shrader-Frechette 2002; Singer 1975; and Sandler and Pezzullo 2007). See also Schmidtz (2001, 2017a).

Among political animals, the fact that people decide for themselves is the basic political fact of life. We can take for granted that people have a common interest in avoiding collision. Beyond that, we also can take for granted that they have destinations of their own. Justice never requires us to pretend otherwise, in part because collisions become lethally inevitable when we do pretend otherwise.

Philosophers are taught to overlook this fact and, instead, grossly exaggerate our need to settle which one of us is right. The truth for political animals is that since we began to settle in large communities, being of one mind has not been an option. Being on the same page is not an option. Our differing ideas about how to resolve conflict are a source of conflict. Our theories do not help. We all have our own theories of value. The fundamental political fact of life is that we live among people who do not find each other's theories compelling, each of them perfectly aware that there is no reason why they should.

Third, *justice is an adaptation*. Ecological justice is an adaptation to reality. It is a response to reality. It is part of how a community responds, and adapts, to problems characteristic of that community's reality. This is not a mere platitude but a consequential working hypothesis. Thus, for example, utilitarian philosophers speak of maximization, but to an economist, maximization is not meaningful except as a response to a set of constraints—which is to say, maximization is not meaningful except as an adaptation to a set of constraints that is subject to change. Likewise, norms of ecological justice are evolving and contingently successful responses to real conflict. They are a society's way of resolving, avoiding, and mitigating conflict. Norms of ecological justice need not be ideals. But to be vindicated as genuine, an alleged ideal has to be a community's ideal response to a real problem—ideal in practice, not only in a theorist's imagination.

Fourth, *justice pertains to process*. For justice to be an adaptation to reality, it has to be an evolving solution to an evolving problem, as opposed to being a response that will be unable to handle what happens next. To me, this suggests that to think ecologically about justice is to think more about processes and procedures than about outcomes. I don't mean to say justice is never about slicing the pie. But I do want to say this: "pie" is a theoretical construct that, at best, represents reality in exactly the way that a photograph does. Although it may be as accurate as a photograph can be, what it represents is a fleeting moment in an ongoing process. Justice is about the process, not the frozen moment represented by the photograph.² So, even if a photograph is as accurate as a photograph can be, it can fail to tell the real story, which is that, in a real ecology or economy, ongoing process is all there is.

A crucial observation is that evolutionary selection processes aim at no particular outcome. Justice need not aim at outcomes either. As I conceive of it, ecological justice is less about how to *distribute* what bakers bring to the table and more about how to *respect* what they bring to the table. Respecting what they bring to the table is a bigger part of making people feel like their contributions (to the process of sustaining a community) are visible.

² Schmitz (2006) sets out a theory of desert with both *compensatory* and *promissory* elements, acknowledging how the idea of deserving a reward differs from the idea of deserving a chance.

Fifth, *justice is somewhat particular*. Mutual understandings about proper ways of avoiding and resolving conflict will be as universal as the conflicts themselves. If a problem is found in all human societies and if there is only one solution, then all societies will converge on seeing that solution as just. By contrast, when problems differ, or when there are many solutions so that a given society needs only to hit upon one among many possibilities, then members of different societies predictably will have different conceptions of justice.

Sixth, *justice is testable*. Ecological justice is not untestable opinion; if it is an ideal, it is not merely a feeling or intuition. Current theories of justice tend to be little more than intuitions. The only way for a philosopher to defend them is by process of elimination: that is, attacking other theories, then imagining for no good reason that one's own pet theory is the one left standing.

There is a kind of ideal theorist who says, in effect, "I am imagining a perfect world. If you wonder what I mean by perfect, the answer is that I am imagining whatever a world would have to be like in order for what I want to call justice to be an ideal response to it. If you want to say the world that I am imagining is unrealistic, I remind you that I am an ideal theorist. I'm not obliged to be realistic. Neither must my ideals be subject to testing and disconfirmation. It is enough that I find them inspiring."

By contrast, I conceive ideals of ecological justice as those that work in actual practice, not only in a cartoon model. Real ideals are not retreats from testability. In different words, an ideal institution (*an institution worth wanting!*) predictably brings out the best in the actual crooked timber of humanity. If an alleged ideal predictably brings out the worst, it is not ideal.³

Science is not geometry. It is a realm of contingent truth, not of necessary truth. Scientific conclusions start from observation, then jump. The same is true of any ethic disciplined by the logics of economics and ecology. We reach conclusions about what ought to be by jumping. The best we can do is to jump from solid information about how norms, rules, principles, and institutions historically affect people's ability to live satisfying lives together.

These features do not define ecological justice, but they are core characteristics of an ecological justice. They indicate whether a given conception of justice is more or less ecological. I have no settled view about how ecological a theory ought to be, but the features listed are ones I would want to see in a theory of justice. Some could be reclassified as corollaries of other features. This list is simply one way of representing the current state of my attempt to clarify the characteristics of a specifically ecological justice.

25.3 BEFORE JUSTICE

To simplify one of the most crucial ideas about the problem to which ecological justice is a response: nature does not obey us. Nature does not have a mind of its own, but does have a logic of its own. If humanity is to make progress, we have to work with

³ See Schmidtz 2017b.

nature's logic rather than against it. Even dictators cannot dictate how things are going to go, any more than exterminators can dictate whether insects will evolve resistance to a pesticide. Every action has more than one effect, and more than its intended effect. Further, and crucially, what you do not see coming will matter. As the poet Robert Burns observed, even the best-laid plans never take everything into account. We plan, but to plan is to focus. To focus is to ignore what is in our peripheral vision. And sooner or later, what we ignore will matter. Overconfidence in a theoretical solution, especially in the light of apparent disconfirmation by the historical record, does not make for justice (especially if learning that an alleged ideal has failed in practice has already driven theorists to retreat from testability to the claim that ideals need not be serious as responses to problems).

Finally, humans are social animals and political animals by nature, and the big-picture ecology of political animals is (among other things) a picture of conflict and negotiation. Other people have hopes and dreams of their own, and their dreams are driving parts of our ecosystem's logic. When we ignore that logic, it comes back to haunt us. We live in a sea of politics in the same way that fish live in water. Being able to anticipate, avoid, and live with conflict is at the heart of a political animal's rational life plan.

25.3.1 Neanderthals

Humanity began to reconstruct its natural world and *Homo sapiens* became the wisest of primates around 40,000 years ago when we learned how to make deals with strangers. In the process, human communities became unlike anything ever seen. Expanding beyond small, face-to-face bands of hunter-gatherers, people began to build far-flung trading networks from which everything we now call civilization arose.⁴

As we spread, humanity's closest cousins, the Neanderthals, went extinct. Why? Theirs was an astounding form of life and had a long run, but *Homo sapiens* had something that *Homo neanderthalensis* lacked. Consider this: 500,000 years ago, Neanderthals formed hierarchical hunter-gatherer groups of about two dozen. Forty thousand years ago, Neanderthals still lived in isolated groups of two dozen hunter-gatherers. It is hard to imagine a human society remaining so static for 460,000 days, let alone 460,000 years. The implied explanation of why Neanderthals (their way of life if not their genes) disappeared and modern humans flourished is that humans were entrepreneurs and Neanderthals were not. Among Neanderthals, there is little evidence of trade. There is little evidence of anything like a tireless search for new ways of being of service. Cultural cross-fertilization did not occur. Neanderthals did not experiment much, and did not learn much from each other's experiments. By contrast, modern humans evidently practiced division of labor from the start, trading services both within groups and between groups. Humans innovate. Humans invent and reinvent their ecosystems.

⁴ See Schmitz and Brennan (2010).

Why? Why did Neanderthals not learn as humans learned, expanding their networks of cooperation as humans did? George Grantham speculates that what distinguished humans from Neanderthals “was not their respective cranial volumes, which overlap, but modern man’s capacity to articulate consonants and vowels required to sound an extensive vocabulary of distinct words” (2008: 8). With an evolving capacity for speech came an ability to make one’s intentions known in environments where effective cooperation was literally life and death.

As the faculty of speech evolved, so could complex trading relationships built on a mutual understanding of such variables as the time and place of one’s next meeting, the exact nature of what was to be delivered, contingency plans, and so on. As potential benefits of trade mushroomed, selection pressure would favor being able to communicate detailed thoughts and intentions. Cooperation increasingly became a matter of communicating with strangers. To sustain expanding networks, people had to develop written language, accounting tools, and a sense of what to expect from each other.

25.3.2 Humans

As trading networks grew, people needed to develop the concept of a legitimate expectation. Thus, a capacity and need for gossip evolved, and with it the possibility of having a reputation. People needed to relay information about who can be trusted—who delivers as promised. From there developed concepts of law and morality. To have business partners 100 miles away, potentially serving thousands of customers, people had to develop the idea of a schedule, the idea of a contract, and the closely related moral idea of a promise.

As the possibility of relying on a partner’s reputation evolves, so does the possibility of vast trading networks. At human sites (but not Neanderthal ones), archeologists find tools made hundreds of miles away. Trade goods traveled long distances, linking language groups and cultures, and contributing massively to the spread of ideas. To see how other groups do things is to see different, and sometimes better, ways of doing things. This very idea—there may be better ideas out there!—may have been the most inspiring of all. Adaptability had become a feature of the phenotype; it had become a species-characteristic survival mechanism. Processes of natural selection could now operate on ideas, not only on genes. It was not only genes that could make copies of themselves. Ideas could make copies of themselves as well. Ideas could go viral.

Humans thus evolved capacities for cooperation that, in degree if not in kind, are unlike anything else under the sun. But along with all the other aspects of human intelligence, humans evolved a capacity for self-awareness, and eventually a conscience. We have become so powerful that we can afford to care about the consequences of using our power in one way rather than another. We can afford to care not only about surviving but about whether we should be proud of what we do to deserve our success.

As our communities evolved, the human condition took shape, and justice took shape with it. It was an evolving aspect of human culture that enabled people to learn what to expect from each other, partly by enabling people to develop ideas about what people have a right to expect from each other. Justice may be more than that, but it is at least that much. It answers the question of how a person ought to live—based in part on what members of a community need to be able to expect from each other if they are to thrive together.

Humanity faces ethical choices. Plato and his student Aristotle, more than anyone else, launched the tradition of Western philosophy when they wrote that we face choices not only about how to survive but how to make life worth living. Two millennia later, Scottish Enlightenment studies of the human condition were the chrysalis from which all of today's social sciences emerged.

Here is a Scottish Enlightenment way of describing permanent features of the human condition. These are features with respect to which justice is an adaptation, a coping mechanism. (Note the ecological dimensions!) First, we make decisions every day about how to handle accessible resources, natural and artificial. Second, resources tend to be scarce: access to resources is costly and limited, and my decisions can interfere with yours. Third, scarcity is manageable if we cooperate (and otherwise generally not). Fourth, people are intelligent; cooperation is possible. None of us can know everything, decide everything, or do everything. Yet, people anticipate and respond to each other. Then, sometimes, we do something gloriously and uniquely human: we give our word, by our deeds we make our word count for something, and in time we learn to trust one another.

Eons ago, people responded to these enduring features by inventing justice: recognizing good reasons for regarding one thing rather than another as what people ought to be able to expect from each other and treat as their due. At the simplest level, justice is traffic management. People need to know what to expect from each other simply to avoid colliding. Drivers need to be able to expect other motorists to stop at red lights. By learning to coordinate on a mutual understanding of whose turn it is, we all end up in a better position to get where we want to go.⁵

We then develop far more elaborate expectations when we trade. Ultimately, we can show up in our community knowing what we want and what to expect from other people, even as we get better at guessing what others want from us in return. We learn how to make people glad to have us around. This is humanity's way of surviving and ultimately prospering.

These are enduring features of the human condition. They motivate us, and enable us, to live mutually respectful and mutually beneficial lives. Our working together to make our world a better place does not end here, but this is where it starts.

⁵ We need not choose traffic management as our problem. There is a testable fact of the matter about whether societies flourish when the sense of entitlement (under specific circumstances, such as when children are involved) encompasses something beyond that.

25.4 BEFORE FAIRNESS, THERE MUST BE TRAFFIC MANAGEMENT

To Bernard Williams, the first question facing a community is about securing “order, protection, safety, trust, and the conditions of cooperation. It is ‘first’ because solving it is the condition of solving, indeed posing, any others” (2005: 3). As communities answered the first question and became more peaceful and cooperative, a second question arose: how to distribute cooperation’s fruits? Philosophers conventionally treat the second as if it were *the* question of justice.

As Williams understood, that second question cannot come first. It is narrow, and does not start at the start. Part of what it means for anything to be a matter of justice is that people have reason to take it seriously. As a general rule, taking it seriously presupposes that taking it seriously is affordable. So, one place to start in developing a theory is by first asking what people can afford.

So far, this picture is illuminating, but not the whole truth. (No picture ever is.) When Williams warns against seeing justice as fundamental, he means justice as a narrow concern about how to distribute an economic pie, as if pie were sitting there waiting to be sliced. Williams is correct that this cannot be the first question. There is a prior question of how societies acknowledge and respect what bakers do to bring the pie to the table in the first place. That prior question needs answering so that people can trust each other enough to risk bringing anything. What we call pie is a metaphor for productivity: productivity is process, not thing; it is flow, not stock. So, Williams is correct to say narrowly distributive justice cannot come first, yet what comes before that static justice (namely, the question of how to respect bakers) is also a question of justice, albeit a more dynamic justice.

Dynamic justice—our segue into ecological justice—responds to these and other features of the human condition. It matters that we arrive in human communities as helpless infants. Moreover, it matters that we do not arrive at the same time, and neither do our claims to the pie. What we call pie is not a thing from nowhere and indeed is not a thing at all. Rather, what we call pie is actually work: the embodied hopes and dreams and lives of other people. So-called pie is an evolving process that exists only because some people keep paying a price to bring services to the rest of us. Therein lies a crucial grain of truth when politicians say, as Elizabeth Warren famously said in a speech in 2011, “Nobody in this country got rich on his own.”⁶ What Warren herself inferred from this premise—that no one owns 100 percent of what they help to produce, therefore they should pay higher taxes—was a non sequitur, but skepticism about Warren’s inference should not blind us to the truth of her premise, or to her discussion’s true implication. The true implication is that no person deserves to be paid more than the value of that

⁶ <https://www.youtube.com/watch?reload=9&v=AHFHznU-N-M>.

person's contribution to a nation's wealth; neither does anyone's contribution deserve less. We respect those whose work created the pie when we treat their contribution as defining their due.⁷ We disrespect those whose work created the pie when we disregard what they contributed and go straight to the question of how to divide what they contributed, as if that were the first question (or even the first question of justice). To think that way is to conceive of the pie as a pirate might conceive it, that is, as loot.

What we all have every right to ask for is a piece of the action, not a piece of the pie: opportunities to learn how to make contributions that will be respected as bases for our claim to fair compensation by identifiable people to whom we provide an actual service. Dynamic justice regards no one as born with a right to command or a duty to obey (Gaus 2000: 143). We are not pawns. The people with whom we do business are our partners, not our owners. The bare fact that no one else owns us—we are each self-owners—is the fundamental equality.

In summary, according to some classroom theories of justice, the order in which people arrive on the scene is the paradigm of a fact that cannot matter, morally speaking. That shows how misleading classroom theories can be. In reality, one key aspect of the human condition is: we are all latecomers. We all face hard questions about how to respect those who came before, and those who will come after. Sustainability is a central focus of dynamic justice.⁸

25.5 CONFLICT AS A FOUNDATION OF COMMUNITY

Constructive legal judgments are adaptations: judges succeed when they settle disputes in a way that helps people stay out of court. Judges aim to honor an individual's right to say no without giving litigious individuals a right to say no to what is best classified as none of their business.

Judges can make a difference. They can ask themselves, "What will people come to expect from each other if I decide the case this way rather than that? What precedent am I setting? Am I establishing a framework for mutual expectation that helps people to

⁷ Of course, those who educate, train, and equip the pie's bakers are also contributing. Each deserves compensation. If we led them to expect that we would pay them, then did not pay them, then we did them an injustice. That too is a legitimate inference from Warren's premise (Schmidtz 2006).

⁸ *Carrying capacity* is among ecology's core concepts, referring to the level of use or traffic that a resource can sustain without degradation. One way to think about justice if we were thinking of justice as an ecological concept would be this: if you want to be fair to future generations, stay within the carrying capacity of your renewable resources. If you do use them up, at least gesture in the direction of justifying such action by providing future generations with something that you think will be more valuable to them than the resources you used up in order to provide them with it (Schmidtz 2000).

get on with their lives in peace? Am I closing a gap between individual and collective interest? Am I leading people to want to make sure their community is better off with them than without them?" Judges who find answers are doing what judges need to do to make their communities better places to live.

Crucially, no one is justly tasked with doing everything worth doing. Instead, in a thriving society, people develop specialties, and it comes to pass that plumbers have a certain job, teachers have a certain job, and so on. Judges in courts of justice also have a job, namely to help litigants resolve conflicts, and derivatively to help people like the litigants to avoid such conflicts in the first place. And as we learned from the Western civilization that emerges from the Reformation, our most reliable as well as most liberal conflict resolution mechanism is to remind litigants that we all have lives to live and to settle who has the right to make the call—who has jurisdiction as recognized by the court—in this particular case. The first imperative, as a method of helping people to avoid or resolve conflict, involves drawing jurisdictions that limit the ability of people to impose external costs on neighbors (that is, costs paid by bystanders not directly consulted in decisions to produce, exchange, or consume any given product).

The second imperative for a civil court judge, arguably, is to limit *transaction costs*, which include all costs of getting goods and services to consumers, including transportation, storage, how long customers wait in line, and so on.⁹ Literal traffic management spells out who owns the right of way at any given moment. Commercial traffic management is a far more complex task, but the basic idea is the same: to help people know what to expect from each other so that they can stay out of each other's way at a minimum, and ultimately see how to help fellow travelers get where they want to go. Giving every landowner a legal right to stop air traffic is a kind of red light that would gridlock traffic, not facilitate it. So, imagine airlines needing to negotiate with every landowner for permission to pass over the owner's land. Air traffic would never get off the ground.

As noted, facts per se do not entail conclusions about the nature of justice; rather, facts reveal the costs and benefits of seeing justice in one way rather than another. Property's purpose as a means of production (how the law has to evolve in order to continue to serve its purpose) has to condition the contours of justice, not the other way around. This brings us back to Bernard Williams's insight: taking justice seriously involves treating justice as something that comes second, not first, because taking justice seriously involves treating justice as something a society can afford to take seriously. It is part of the larger context of property rights that the right to say no is not a weapon of mass destruction. It is a right to say no to being involved in a transaction, not a right to gridlock the system by saying no to other people transacting.

Judges aim to resolve disputes and leave litigants in a position to get on with their lives and livelihoods. So judges must aim to send litigants away with stable and robust mutual expectations. The judges also needed have one eye on settling a dispute in such

⁹ See Schmidtz (2011) for a discussion of landmark cases in the evolution of our thinking about the twin legal imperatives to minimize transaction cost and external cost.

a way as to avoid having the same dispute come up repeatedly. Judges aim to settle a type of dispute, not merely a particular instance, so as to enable people in general to understand how to stay out of each other's way, stay out of trouble, and stay out of court. Judges tend not to be striving for what academics call justice. They strive for the kind of justice embodied in a peaceful community, not the kind embodied in a classroom theory.

25.6 VARIETIES OF ENVIRONMENTAL CONFLICT

Think of environmental conflicts as disputes where environmental impact is part of the problem. Fifteen years ago, I described three varieties of environmental conflict. Part of the plan was to plead for philosophy that is alive to social science in general and economic analysis in particular. While reducing all values to economic values is uncalled for, this key fact remains: economies exhibit a certain logic. To ignore the logic of human economy is to ignore the logic of human ecology and thus to ignore the logic of ecologies where humans are a keystone species. What a shame it would be to let ostensible environmental values lead us to such a crippling obliviousness to ecological reality. Conflict occurs when people get in each other's way. Think of a traffic jam as a problem that results when people get in each other's way. In particular, when the uses to which different parties would put common resources or common spaces leads to congestion, the effect is that people aiming to use such resources find themselves in each other's way. Societies cope by developing institutions that (in effect) manage traffic. The tools of literal traffic management—stop lights and speed limits—are paradigms.¹⁰

Notice that if we look at a snapshot of a system of green and red lights, we will appear to be looking at a zero-sum game. An observer who sees only the snapshot will appear to see the rich getting richer and the poor getting poorer. The rich (those whose light is green) get to move while the poor (those whose light is red) are excluded, stopped, kept in place. We take offense, but the inequality that offends us is snapshot inequality. Snapshot inequality is a mirage, and the snapshot perspective is a fool's perspective. The snapshot fails to represent what matters: namely, this ecology is in motion. It is a process, not an outcome. Even if everyone has to wait for their turn, the snapshot makes it appear that life is a green light for some people and a red light for others. It matters whether everyone has a chance to seek their own destination; what matters not at all is whether

¹⁰ Schmitz (2000) describes this as conflict in use, contrasting it with two other types of conflict. A second is *conflict in values*, such as when people disagree not about whose turn it is but whether anyone should get a turn at all (for example, to hunt elephants). A third conflict is *conflict in priorities*. Subsistence farmers for whom getting enough food is a day-by-day proposition can have priorities unlike ours not because their values are different but precisely because their values are the same. This conflict is not a difference of values but a difference in which values people can afford to pursue under different circumstances.

everyone gets to go at the same time. If everyone (including pedestrians) gets a turn eventually, then the system is egalitarian in the exact way a system needs to be egalitarian in order for people to be able to live good lives together.

Property rights are more complex than traffic lights, but the same point applies. Norms of ownership establish who has the right of way, who must say yes for exchange to be legitimate and who has a right to say no. Property rights manage the traffic of beings who survive by trading. Property institutions help people avoid, manage, and resolve conflict when they (1) facilitate orderly use of a common resource, (2) facilitate orderly removal of resources from the commons, and (3) help people cope with externalities, including new externalities that emerge as property regimes evolve.¹¹ Property regimes can solve commons problems and induce overall patterns of sustainable use. Allocating rights to enforce a rule of “no trespassing” or “no hunting” is a method of managing access to and thus traffic on a given parcel of land.

No system of property will look fair if all we see is a snapshot. It is not possible. Twenty-year-olds always look like grossly disadvantaged have-nots compared to forty-year-olds (unless an economy totally collapses). Then, as a society gets richer, that increasing productivity manifests itself as forty-year-olds pulling ever farther ahead of twenty-year-olds, by virtue of having twenty more years in an increasingly productive society. Still, current snapshot measures of income inequality show us nothing but a simplistic picture of rising inequality. If society is getting richer, then the twenty-year-olds ultimately will have higher lifetime income than current forty-year-olds, but in the snapshot, such dynamics will be invisible. Meanwhile, the serious question will always concern what a society makes possible over a lifetime.

Environmental philosophers have been silent on environmental conflict resolution, and have focused instead on environmental justice. But the road to a true understanding of environmental justice passes through the terrain of environmental conflict. That is why there is a general unease among consumers of this literature: there is something “ivory tower” about what we call environmental justice. The rhetoric too often is a form of posturing not even intended to inspire conflicting or potentially conflicting parties to work toward building better lives together.

When justice is presumed to dictate destinations rather than manage traffic, it fails to coordinate. It fails to solve a problem. Progress comes to a halt. Actual progress starting from here needs the cooperation of stakeholders who have destinations unlike ours—stakeholders who have no idea why we imagine that our destination is more important than theirs. The justice that enables people to make progress together (despite having different destinations!) will be unlike the conceptions of justice that we dream up in classrooms.

¹¹ Note: we aim not to eliminate external cost so much as to keep the cost within reasonable limits. Driving has external costs, but we do not want people to stop driving. Being a good neighbor will always involve tolerating minor irritations. It also will involve taking reasonable steps to avoid being an irritant. There is no substitute for being considerate. No system of law will enable us to be good neighbors just by obeying the law.

When we say our principles should be put into practice, we imply that our principles are consistent with sound practice. We imply that we have done our homework. If we have not—if we never go beyond toy thought experiments and do the hard work of grounding our theories in requirements of sound practice—then when practitioners ignore us, they are doing the right thing.

25.7 MORAL IDEALS AND POLITICAL COMPROMISE

Philosophers, one could argue, earn a living by envisioning a more perfect world. But although there may be some point where environmental ethics is an exercise in envisioning ideals, environmental conflict resolution is an exercise in the art of compromise. Successful conflict mediation typically involves negotiating win-win solutions. Does that presuppose that both parties would win in a morally ideal world? No. The realistic ideal is to be able to get along with all of our neighbors, not only the ideal ones.

It is a fantasy, not a real ideal, to say that ideally we would not need to get along with nonideal neighbors, because we are imagining a society in which nonideal neighbors do not exist. Realistically, there are times for attending not only to our values but to other people's too, even if we do not share their values, and indeed find them alien. Why? Because other people's values are integral parts of ecosystems about which we claim to care. We do not decide how people will act any more than we decide how elephants or viruses will act. People decide for themselves. If we aim to respond in an environmentally sound way, we need to ask, under what conditions do people with *their* values and *their* priorities act in environmentally sound ways?

So, if we go to the table with an environmentalist agenda, we must understand this: for our stand to be politically effective, our opposition must go home feeling like they were heard and respected. Further, they got something out of the deal; it was not only they who made concessions. Politically, aiming for "win-win" goes with being serious about our moral ideals.

25.8 CONCLUSION: TOWARD REAL WORLD ETHICS

Some of our questions are simply invitations to reflect. What kind of world do we live in? What sorts of love and respect for nature, community, and self are ingredients of a thriving life? Human justice has its origin in the need for cooperation among political animals: animals who survive by negotiating mutually acceptable terms of

engagement. Does ecological justice likewise trace its origins to a need for cooperation, or does it start elsewhere? If it starts there, does it stay there, or does it evolve into something different? What if the human condition were understood as a robust and resilient yet still evolving ecological niche? It is hard to imagine everything that such a question might imply.

First, however, it implies that we live in evolving ecosystems, and cannot thrive except by adapting to a changing world. We cannot depend on our ecological niche to be static, but on the other hand, our niche does have a history of being stable enough that some ways of adapting to it have been robustly better than others. Our ecological niche is a sum of moving parts, always shifting. Our ability to adapt has made us as successful a species as we are.

Second, humanity's ability to adapt is phenotypic more than genetic. Humanity does not wait for mutations to appear in its gene pool, randomly affecting genetic fitness for better or worse, then letting natural selection sort things out over eons. Humanity's survival mechanism lies in our individual ability to adapt on the fly to new and rapidly shifting niches, inventing new niches as we go: buggy whip manufacturer, key punch operator, Angry Birds developer, Lasik surgeon (Millgram 2015). Insects and viruses are highly adaptive, too, but their ability to adapt is genotypic, thus altogether unlike humanity's.

Further, even leaving aside ecological questions and thinking narrowly about justice between humans, we are already talking about a subject where the context matters. If I am asking you what Sue deserves, you need to know whether Sue is an infant (where the question may be best interpreted as a question about what Sue needs) or a colleague standing for promotion (in which case talking about what Sue needs would be patronizing, whereas respecting Sue as an adult colleague requires that we focus on what she has accomplished). Therefore, before we apply traditional theorizing about justice to environmental questions, we are already in a realm where we have to be cautious about extending a theory beyond the subject matter that defines its natural range of application and interpretation. Ecological justice may be a domain whose principles are not like principles that apply in other domains.

Can we start over? Can we reimagine the human condition as an ecological problem? Perhaps environmental ethics deserves better than to be treated as "applied" ethics. Perhaps we can go beyond seeing environmental justice as pretending that traditional theories illuminate dynamic ecological problems and instead start fresh—with ecology.

ACKNOWLEDGMENTS

This essay is substantially new work, and still very much work in progress, yet it also is work that is pulling together everything I've written since graduate school. It is proximately a continuation and reworking of Schmidtz (2016b, 2018). However, I also lean more or less heavily on Schmidtz (2000, 2001, 2006, 2011, 2016a, 2017a, 2017b) plus Schmidtz and Brennan (2010) and Schmidtz and Willott (2003).

REFERENCES

- Bullard, Robert. 1990. *Dumping in Dixie: Race, Class, and Environmental Quality*. Boulder: Westview, 1990.
- Gaus, Gerald F. 2000. *Political Concepts and Political Theories*. Boulder: Westview.
- Grantham, George. 2008. "The Prehistoric Origins of European Economic Integration." McGill University Economics Department Working Papers.
- Hill, Thomas E., Jr. 1983. "Ideals of Human Excellence and Preserving Natural Environments." *Environmental Ethics* 5: 211–24.
- Millgram, Elijah. 2015. *The Great Endarkenment*. New York: Oxford University Press.
- Norton, Bryan. 1991. *Toward Unity among Environmentalists*. New York: Oxford University Press.
- Sandler, Ronald, and Phaedra Pezzullo. 2007. *Environmental Justice and Environmentalism*. Cambridge, MA: MIT Press.
- Schmidtz, David. 2000. "Natural Enemies: An Anatomy of Environmental Conflict." *Environmental Ethics* 22: 397–408.
- Schmidtz, David. 2001. "A Place for Cost-Benefit Analysis." *Philosophical Issues* 11: 148–71.
- Schmidtz, David. 2006. *Elements of Justice*. New York: Cambridge University Press.
- Schmidtz, David. 2011. "Property." In George Klosko (ed.), *The Oxford Handbook of the History of Political Philosophy* (Oxford: Oxford University Press), pp. 599–610.
- Schmidtz, David. 2015. "Corruption." In Subramanian Rangan (ed.), *Performance and Progress: Essays on Capitalism, Business, and Society* (Oxford: Oxford University Press), pp. 49–64.
- Schmidtz, David. 2016a. "After Solipsism." *Oxford Studies in Normative Ethics* 6: 145–65.
- Schmidtz, David. 2016b. "Ecological Justice." In David Schmidtz (ed.), *MacMillan Handbook of Environmental Ethics* (Boston: Cengage Learning), pp. 221–42.
- Schmidtz, David. 2017a. "Environmental Conflict." In Stephen Gardiner and Allen Thompson (eds.), *The Oxford Handbook of Environmental Ethics* (New York: Oxford University Press), pp. 517–27.
- Schmidtz, David. 2017b. "Realistic Idealism." In Adrian Blau (ed.), *Methods in Analytical Political Theory* (Cambridge: Cambridge University Press), pp. 131–52.
- Schmidtz, David. 2018. "Public Choice as Political Philosophy." *Review of Austrian Economics* 31: 169–76.
- Schmidtz, David, and Jason Brennan. 2010. *A Brief History of Liberty*. New York: Blackwell.
- Schmidtz, David, and Elizabeth Willott. 2003. "Reinventing the Commons: An African Case Study." *University of California Davis Law Review* 37: 203–32.
- Shrader-Frechette, Kristin. 2002. *Environmental Justice Creating Equality, Reclaiming Democracy*. New York: Oxford University Press.
- Singer, Peter. 1975. *Animal Liberation*. New York: Harper Collins.
- Stone, Christopher. 1996. *Should Trees Have Standing?* Dobbs Ferry, NY: Oceana.
- Williams, Bernard. 2005. *In the Beginning Was the Deed*. Princeton, NJ: Princeton University Press.

CHAPTER 26

CIVIL RIGHTS, EMPLOYMENT, AND RACE

BRENDAN O'FLAHERTY

26.1 INTRODUCTION

ALMOST all economists support civil rights laws, especially those for employment, but most find it hard to explain why. What is the market failure that these interventions are supposed to be addressing? Or are they compensating for slavery and Jim Crow? Are they supposed to improve income distribution?¹ Should they do nothing more than penalize failures to maximize profit—and if so, isn't that what markets do? In a situation where they did not exist, how could you tell whether they were needed—and in a situation where they already exist, how could you tell whether they were still needed? Why do they cover certain decisions, such as whether to hire someone, but not others, such as whether to work at a particular firm? Are there better ways of accomplishing whatever it is that civil rights laws are supposed to be accomplishing?

Here I concentrate on the United States, on race, and on employment: the United States, because I live there; race, because it was the original subject of these laws in the 1860s and 1960s; and employment, because it is the main source of most households' income and a major contributor to many individuals' sense of self-worth. But I will look beyond these confines as needed for examples and contrasts, particularly in the first half of the chapter.

The plan of the chapter is to begin with theory, mainly from philosophy because economists have made few attempts to develop a theory of civil rights laws. This part is

¹ The contribution of racial disparities to overall income inequality is in fact tiny. If the household income distributions of all racial and ethnic groups were the same as the income distribution of non-Hispanic whites in 2016, the Gini coefficient would fall only from 0.481 to 0.469 (from table H-4 in US Bureau of the Census 2018). It is trivial to calculate what would happen to the aggregate Gini if all racial groups had the same income distribution as NHWs do: the aggregate Gini then would be the same as the NHW Gini.

normative. Then I will try to assess what economists have learned about employment in the United States and about civil rights laws in terms of what the theory says about the environments in which civil rights laws should operate and what effects civil rights laws should have. This part is positive. Finally, I will speculate about the future, both in an academic sense, in terms of what questions should researchers try to answer, and in a policy sense, in terms of what modifications, if any, should be made to these laws in the twenty-first century.

26.2 WHAT IS THE PURPOSE OF CIVIL RIGHTS LAWS?

Employment policy is part of civil rights legislation. The primary law on employment discrimination, Title VII, was part of the Civil Rights Act of 1964. Civil rights can be defined as “a cluster of individual entitlements that protect individual dignity and collectively constitute basic, minimum requirements for free and equal membership in democratic society” (Shin 2013: 794). They are different from human rights or legal rights generally, but no consensus exists on what makes them different or what the precise boundaries are. A person’s “having a right to x entails the impermissibility, at least *prima facie*, of depriving that person of x , or that such deprivation can only be justified by strong reasons that go beyond a simple utilitarian balancing of costs and benefits” (2013). Some civil rights restrict what governments can do (such as voting rights) while other civil rights restrict what private citizens can do (public accommodations and employment in particular). We concentrate on the latter.

26.2.1 Justice and Rights

Most of the philosophers who write about race normatively do not address policy in the utilitarian framework that economists use; instead they concentrate on justice and rights. Rawls expresses the view this way:

Each person possesses an inviolability founded on justice that even the welfare of society as a whole cannot override. For this reason justice denies that the loss of freedom for some is made right by a greater good shared by others. It does not allow that the sacrifices imposed on a few are outweighed by the larger sum of advantages enjoyed by many. Therefore in a just society the liberties of equal citizenship are taken as settled; the rights secured by justice are not subject to political bargaining or to the calculus of social interests. . . . [A]n injustice is tolerable only when it is necessary to avoid an even greater injustice. (1971: 3–4)

The policies that analysts in this tradition advocate, then, do not necessarily maximize or minimize anything. Instead, they assure that everyone has the “liberties of equal

citizenship” that are appropriate. These analysts may differ from Rawls on almost everything, but on the primacy of rights they agree.

Because Rawls and these analysts are not utilitarians, they can dismiss the benefits of racism to racists. Utilitarians are often embarrassed to count the benefits of racism as regular benefits (just as they are embarrassed to count the benefits of crime to criminals). Rawls, on the other hand, finds no problem in dismissing these benefits:

In justice as fairness... persons accept in advance a principle of equal liberty and they do this without knowledge of their particular ends. They implicitly agree, therefore, to conform their conceptions of the good to what the principles of justice require, or at least not to press claims which directly violate them. An individual who finds that he enjoys seeing others in positions of lesser liberty understands that he has no claim whatever to this enjoyment. The pleasure that he takes in others' deprivations is wrong in itself... The principles of right, and so of justice, put limits on what satisfactions have value; they impose restrictions on what are reasonable conceptions of one's good. (1971: 31)

26.2.2 How Does Employment Discrimination Violate Someone's Rights?

To economists, the form of civil rights laws looks strange. In many cases, the refusal of person A to purchase something from person B makes B worse off, but it is usually not argued that B's rights are thereby violated. Some people never eat mashed potatoes, and potato farmers are worse off as a result. But no one claims that this refusal to eat mashed potatoes is in general a violation of potato farmers' rights.

Several rights-based answers to this challenge involve particular aspects of the discriminatory process that seem to make it socially worse than simple refusals to purchase labor. Mental-state-based accounts say that discrimination shows a discriminator's lack of respect for the person discriminated against, which is bad in itself and particularly stinging to the latter (Lippert-Rasmussen 2013). Objective-meaning accounts find the wrongness of discrimination in its demeaning nature, independent of the moral state of the discriminator. Both approaches agree that discrimination is bad in a particular way for the person discriminated against.

The third class of answers are called “harm-based,” but they take a wider view of social interaction and tie to the idea of civil rights more closely. Examples are Lippert-Rasmussen (2006) and Moreau (2010); I will concentrate on the latter. Moreau starts with the following position:

In a liberal society, each person is entitled to decide for herself what she values and how she is going to live in light of these values... This means that, in addition to certain freedoms of action, we are each entitled to a set of ‘deliberative freedoms,’ freedoms to deliberate about and decide how to live in a way that is insulated from pressures stemming from extraneous traits of ours. (2010: 147)

The job of civil rights law, then, is to assure each person of her “individual entitlement” of deliberative freedom.

The extraneous traits involved need not be either innate or unimportant to the person who is deliberating—being black or a woman or a Muslim may be core to her identity (or not). But the right to decide *how* being black or a woman or a Muslim affects her deliberation or life is hers, not an employer's, bar owner's or landlord's. According to Moreau, a religious person

may, for instance, rule out certain places of work and certain neighborhoods on the grounds that he would encounter too many people who dressed immoderately. Yet we still believe his religion to be normatively extraneous to these deliberations in the normative sense that the social costs of practicing it are not costs that we think should have to figure into his deliberations. (2010: 149–50)

The right is quite specific. Employers can judge other job-related competence, and prospective spouses can judge religion. And the right is everyone's. So in this sense it is a universal minimum entitlement in the space of rights. Why the right should be a universal minimum is obvious—it is a prerequisite to flourishing.

The right, however, does not always apply in employment situations where it conflicts with other rights. Moreau gives the example of “bona fide requirements”: anti-discrimination law does not require the Catholic Church to hire Muslims as priests, or theatrical directors to hire women to play Hamlet. Moreau presents these as *conflicting rights*—the right of the director to interpret *Hamlet* in her own artistic terms versus the right of the highly qualified actress to play Hamlet.

Moreau uses the conflict of rights to explain why anti-discrimination law is confined to areas like public accommodation, employment, and rental housing, and does not apply to areas like choosing marriage partners, making friends, rooting for baseball teams, or voting for political candidates:

In my view, we can see the limitations on social contexts in which anti-discrimination law operates as attempts to foster certain values other than freedom. One of these may be the value of deep personal relationships. One plausible reason why anti-discrimination law does not regulate personal dealings between family members and friends is that, for caring relationships to flourish, people need to be able to choose whom they wish to be close to and how they are going to treat each other and they need to be able to treat their loved ones in special ways, ways in which they do not treat every other person. (2010: 161)

Moreau's perspective of anti-discrimination as freedom makes the clustering of civil rights laws more obvious. Being hired is not much use if one can be immediately fired, and protection against firing is not much use if one is never hired. One cannot pursue the occupation of one's choice unless one can get the education that that occupation requires, be allowed to meet in restaurants and hotels with the people the occupation requires one to meet with, and not get shot by the police or incarcerated. Similarly, living

in a safe pleasant neighborhood or sending one's kids to a great school is hard unless one has a reasonably secure job that pays well. The various civil rights are complements. It is more than a coincidence that in the United States almost all the relevant legislation was enacted at around the same time, and often in the same bill.

Complementarity may also operate within and between firms. In a society with a history of racism, within-race interactions and relationships may be less fraught, more comfortable, and more productive than between-race interactions and relationships. For many workers, interactions are extremely important for output—either within the firm (with their supervisors), with other firms (with purchasers), or with customers (with sales people). An employer contemplating hiring a black worker, therefore, will be more likely to do so if she knows that there are more black workers in the firm already, believes that she will hire more black workers in the not-too-distant future, believes that other firms will hire more black workers, or believes that more blacks will be hired generally and become richer and better customers. Civil rights laws, because they are applied generally, shift all employer expectations in the direction of greater productivity for prospective black workers, and so encourage their hiring. Laws that cover all employers can achieve coordination that individual bargaining cannot.

All of the rights-based answers help answer a question that often occurs to economists and economics students: why do civil rights laws prohibit discrimination when doing so forces innocent non-racist people to suffer objective losses? Examples are restaurant owners who might lose customers if they hired African American waitstaff and police officers who might recover less contraband if they stopped white drivers more often (“strategic discriminators”). The answer from rights-based analysts is simple: violating someone's rights for objective gain is wrong because it's a violation of rights, and just as wrong as violating someone's rights for racist reasons. Bernie Madoff was just as much a criminal as Jeff Dahmer. Just as a police officer who engages in racial profiling to increase contraband recovery is more understandable than one who profiles because of racial animosity, Madoff's greed is more understandable than Dahmer's mania, but both of them are criminals.

26.2.3 When Are Laws Needed?

Laws designed to guarantee deliberative rights can also come into conflict with other rights in ways that cannot be legislated around—for instance, employers' rights to hire people and therefore associate with them for reasons that are unconventional, trivial, or silly—and they also can impose large costs in a consequential sense. The reason is that enforcement is expensive, and often, because enforcement relies on a counterfactual, faulty. Courts will make mistakes, and even if they do not are expensive to deal with. Moreover, these laws can give employers incentives to evade them—by moving away from places where minorities are likely to find them, for instance, or developing elaborate procedures to make it appear that they are complying when they are not. These too are costly. So there are prudential reasons not to use laws to guarantee deliberative rights

to people who would have such rights in the absence of laws, or for whom such laws (if perfectly enforced) would make a trivial difference.

Laws make a much greater difference to the deliberative freedom of prospective employees who face pervasive discrimination and may make a negligible difference for those who do not. A prospective employee whose extraneous traits are objectionable to no employers gains nothing from the law; and one whose extraneous traits are objectionable to only a few employers also gains little or nothing. With very good information, at any moment a prospective employee benefits from a law only if at least one of her best two matches in the labor market finds her extraneous traits objectionable. Because of imperfect information flows and ever-changing conditions, however, we can say roughly that only large-scale discrimination makes laws assigning deliberative rights to potential employees produce large benefits.

The expressive harms of discrimination may work in the same direction, although this conclusion is more controversial (see Risse and Zeckhauser 2004 and Durlauf 2006 for contrasting positions). Consider Claude Steele's account of his reaction to learning as a child that he could not use the municipal swimming pool in Chicago because of his race:

To my seven- or eight- year old self this was a bad condition of life. But the condition itself wasn't the worst of it. For example, had my parents imposed it on me for not taking out the garbage, I wouldn't have been so upset. What got me was that it was imposed on me because I was black. There was nothing I could do about that, and *if being black was enough to restrict my swimming then what else would happen because of it?* (2010: 2, emphasis added)

By contrast, when I was in college during the Vietnam War, my birthday led me to be assigned draft lottery number 22. I was unhappy and considered myself quite unlucky, but it never occurred to me that this event might presage a lifetime of oppression (and it did not). Consideration of both objective and expressive harm suggests some empirical measures for assessing how beneficial assigning deliberative rights to different groups is.

26.2.4 What Activities Should Civil Rights Law Cover?

Thus we have a fairly coherent argument for deliberative rights in hiring under certain conditions, and we have some idea of how to test whether those conditions in hold at various times and places. The arguments about hiring extend fairly easily to employment retention, promotion, and working conditions, which are also covered by Title VII. Once workers are hired, employers enjoy some degree of monopsony power over them, and so arguments for assigning deliberative rights to workers are generally stronger than the arguments for potential employees who have not been hired yet. But workers who face less discrimination outside the firm still have better outside options, and so gain less from the assignment of deliberative rights, everything else being equal.

Arguments very similar to the arguments for hiring can also be used for public accommodations, education, and health care, which are also covered by civil rights laws. But many areas that would seem to require coverage do not have it.

Most obviously, why are workers allowed to discriminate against employers, but employers not allowed to discriminate against workers? An aspiring entrepreneur who cannot hire the right workforce because they find her extraneous traits objectionable surely has her deliberative freedom curtailed. There are also no laws against discrimination in choosing which neighborhoods to search for housing, which baby to adopt, or which potential spouse to marry. Moreau, as we have seen, makes an argument against marriage civil rights laws by appealing to the value of deep personal relationships, and this argument may extend to adoption (although it is not clear that babies differ so greatly in dimensions beyond gender and race that selection is needed for a deep personal relationship). But workers and employers do not usually have the sort of deep personal relationship that Moreau appeals to—and if they did, her argument for civil rights laws covering the actions of employers would fail—and neighborhoods probably do not qualify as deep personal relationships either.

There are several other possible explanations for these legal gaps, most of which are tied to enforcement. At the most basic level, enforcement of a civil rights law depends on answering a counterfactual question: what would the worker or the homebuyer or the aspiring adoptive parent or the potential spouse have done if the employer, neighborhood, child, or suitor had some other extraneous trait? Possibly legislators believe the counterfactual hiring question is easier for courts to answer than the counterfactual in the case of marriage, adoptions, or home purchases. It would be interesting to know whether this presumed belief is accurate, especially for highly skilled jobs.

Hiring also requires potential employees to send a costly and hence credible signal that they are probably willing to work for an employer: they must apply for the job (which was more costly in 1964, long before the Internet, than today). By contrast, homeowners selling houses, landlords renting apartments, and children awaiting adoption do not usually send specific costly signals to a small number of homebuyers, apartment-seekers, or aspiring adoptive parents indicating their interest in engaging in a transaction with them. Thus the parties who have been discriminated against cannot identify a unique counter-party who discriminated against them.

This problem of identification of a unique counter-party may also apply to discrimination by workers against start-up firms, but not to discrimination by workers against continuing firms. If a black entrepreneur buys an existing grocery store, makes no change in the conditions of employment, and all the workers there quit immediately, she has specific counter-parties to sue, but no law to sue them under.

The locus of decision-making may also matter: decisions that involve organizations are more likely to be covered than decisions by individuals. Individuals can make decisions without a paper trail, discussions, or help from others; organizations presume some coordination and delegation, which usually are accompanied by communication. Laws may only move organizations from overt discrimination to tacit discrimination, but tacit discrimination may be less effective than overt. So laws governing organizations

may have an effect even when court cases are virtually impossible, but laws governing individuals would have no effect.

The difficulty of enforcement makes a case for civil rights laws to cover every possible area. Because some agents are strategic discriminators and enforcement in no area is perfect, civil rights laws in different areas are, in general, complements to each other.² If, for instance, blacks on average received better education, strategically discriminatory employers who use stereotypes would be more inclined to hire them. If blacks had better jobs, and housing discrimination were less, their children would live in districts that provided better schools. Because enforcement in any one area is likely to have diminishing marginal returns, efficient civil rights strategies include civil rights laws for as many areas as possible (realizing that not all areas are possible).

26.3 WHAT DO ECONOMISTS KNOW ABOUT RACE IN THE LABOR MARKET?

With this background, we can try to see whether the conditions for enacting and maintaining civil rights laws hold—whether some people lack deliberative rights—and whether civil rights laws do a good job of guaranteeing deliberative rights for everyone. Economists have been trying to answer these questions for decades, although they do not describe what they are doing in these terms. Generally, the findings indicate that blacks and Hispanics lack deliberative rights, but whites do not, supporting the case that civil rights laws covering blacks and Hispanics (but not whites) are justified. However, the evidence on whether civil rights laws do what they are supposed to do—assure deliberative rights—is more mixed.

In the rest of the chapter, I will follow the conventions of the US government in defining races, and the conventions of *The New York Times* on capitalization. However, I will also follow common usage by considering non-Hispanic whites (NHWs), rather than all whites, when the data are available. For the most part, I will try to use the partition of groups as Asian, black, Hispanic, Native American, and non-Hispanic white, with black meaning non-Hispanic black. I take this as closest to the way educated Americans partition the American population these days. For instance, the database maintained by *The Guardian* (2017) on victims of police shootings uses these groups (with the understanding, I think, that white means non-Hispanic white), and for most of the twenty-first century the New York City street count of homeless people used it. Philosophers have developed a clear definition of ethnicity as something different from race (Kitcher 1999; Taylor 2013: 52–5), a distinction that is very helpful for clear thinking, but is not reflected in how most people collect and interpret data.

² This is a corollary to Myrdal (1944: Appendix C).

26.3.1 Labor Markets in General

Economists have studied race in the labor market longer and more intensively than in any other area. However, labor markets have not drawn economists' attention as much in recent years as they did before the Great Recession.

In the United States today, NHWs and Asians are more likely to have jobs than Hispanics, and Hispanics are more likely to have jobs than blacks and Native Americans. Among people with earnings in civilian non-institutional population, the ranking for earnings is about the same, but Hispanics do not do better than blacks and Asians do considerably better than NHWs (US Bureau of the Census 2017).

The differences are not trivial. Among adults in the civilian non-institutionalized population, Asians are around a quarter more likely to be employed than Native Americans, and a seventh more likely to be employed than blacks. Among employed adults, Asians earn two-thirds more than Hispanics, blacks, and Native Americans, and NHWs earn half as much more. Combining these two differences, earnings per adult are almost twice as high for Asians as for blacks and Hispanics, more than twice as much as for Native Americans. These data exclude the "group quarters" population, which includes nursing home residents, prisoners and inmates, homeless people, and some college students.

More directly, among employed workers, blacks on average are considerably less likely to be satisfied with their jobs than whites are. This gap has persisted for over four decades, and is not diminishing (Mukerjee 2014; Autor and Kitcher 2018). Standard correlates, including income, cannot fully explain the gap (Mukerjee 2014). Blacks and Hispanics are also more likely than whites to be killed or injured on the job (Leeth and Ruser 2006; Viscusi 2003).

26.3.2 Economists' Questions and Answers about Labor Markets

These gross differences by themselves tell economists almost nothing about whether labor markets are operating in any sense unjustly, unfairly, or inefficiently. The differences could be due to age structure, region of residence, or tastes for work and leisure. Or, labor markets could be working just fine and the disparities in labor market outcomes could be due to the functioning of some other part of the economy: certain groups may have worse educational opportunities currently, or have had worse educational opportunities in past generations that have been handed down to the current generation, or may suffer from policing practices that make it harder for them to get and keep jobs.

Trying to explain these disparities and locate their sources has been the main focus of economists' attention. There is a large body of research, mainly empirical, and the consensus is that non-racial variables cannot explain all of the differences among racial

groups, particularly the poor outcomes of black men, but some controversy persists over the size of what non-racial variables cannot explain. "The difference that non-racial variables cannot explain" is what economists call "discrimination." The idea is that if, for instance, a black man and a white man are identical in all relevant respects, but have different wages or probabilities of employment, then something inexplicable by normal economic causes is going on, and therefore policies to alleviate it are justified. This is essentially Moreau's question about whether both have deliberative rights, which grants these economic results philosophical significance.

The discrimination that economists search for can be divided into two types, depending on how the search is conducted: disparate treatment and disparate impact (O'Flaherty 2015, appropriating legal terms and extending their meanings). *Disparate treatment discrimination* occurs when an individual A, usually an employer in labor discussions, treats another individual, B, worse than A would have treated B if B's race were different. (For some purposes, disparate treatment may also mean a disposition on the part of A to act in this manner, but uses are not always clear.) Disparate treatment discrimination may or may not make individual B worse off. *Disparate impact discrimination* occurs when something in the labor market makes an individual worse off than she would be if her race were different, and everything else about her were the same. The cause, for example, could be a policy or law, individual B's beliefs about how she would be treated, or even disparate treatment discrimination.

Disparate treatment discrimination can occur without disparate impact, and vice versa. For instance, if a small number of otherwise identical employers discriminate against workers of a particular race, but a number of employers large enough to easily employ all members of that race do not discriminate, then disparate treatment can occur without disparate impact. (I may be prejudiced against the Kardashian family, and would never patronize a pizzeria where one of them worked, but my prejudice does them no harm.) Or if there are two races and every employer discriminates against one race or the other, but the volume and virulence of discrimination is the same against both races, then no disparate impact may occur despite rampant and pervasive disparate treatment discrimination. Gary Becker's famous book *The Economics of Discrimination* (1957) is essentially an analysis of when and how disparate treatment in fact leads to disparate impact.

In the other direction, suppose members of some race believe that they will never be hired in a particular job, and so they never acquire the training the job requires. If their belief is wrong, then there is disparate impact discrimination but no disparate treatment discrimination. Or suppose a law forbids a certain group from acquiring the training for the job. The prohibited group is not defined by using explicitly racial terms, but its membership is in fact drawn disproportionately from a particular racial group (in the nineteenth century and first half of the twentieth century, people whose ancestors were not eligible to vote in 1866, or in the twenty-first century, people with connections to nations on the terrorist watch list). Then disparate impact discrimination occurs without disparate treatment.

Generally, economists seem to think that disparate impact discrimination is a more serious problem than disparate treatment discrimination, since disparate impact by

definition must affect someone's well-being. This ordering follows from the profession's utilitarianism. The priority of disparate impact can be seen, for instance, in Heckman's (1998) criticism of employment audits: he stated, in the language I use, that these audits measured disparate treatment discrimination, not disparate impact discrimination, and let that simple statement stand without further elaboration as a strong criticism.

For examining deliberative rights, disparate impact is also probably the more important measure. White men who are or would be denied jobs in barber shops that cater to blacks, but end up at Goldman Sachs, have probably not had their deliberative rights seriously infringed.

26.3.2.1 *Disparate Treatment Results*

To search for disparate treatment discrimination, economists (and sociologists too) generally use some form of "employment audit." In the traditional version of these audits, researchers employ teams of "auditors" from various racial backgrounds. They form pairs of one NHW and one member of a minority group, but try keep the members of the pair as similar as possible in easily observable non-racial characteristics such as age, gender, and height. The researchers then randomly draw employment advertisements, and to each opening they assign a pair of auditors, instruct them on what information to use on their applications—similar but not identical—and randomly assign them times to apply for the jobs—close but, again, not identical. They send the auditors off and record what happens and whether they receive job offers. Of course, the auditors who receive job offers never accept them because they are not really looking for jobs. (Audits involve deception, and so they cannot be used where deception is illegal—in mortgage applications, for instance.)

Each advertised job generates one of four outcomes: both auditors get offers, neither gets an offer, the NHW gets an offer and the minority auditor does not, or the minority auditor gets an offer and the NHW does not. Generally, audits include some of each outcome, with neither getting an offer being the most common. Even if no employer is discriminatory, random variation should result in some situations where one auditor gets an offer and the other does not: sometimes the job is actually filled between the two visits, the employer is in a lousy mood for one visit and a better mood for the other, one of the auditors was unprepared or not feeling well, or someone was delayed on the subway. Even if every employer is discriminatory, some of them will treat both auditors the same because both are either manifestly qualified or manifestly unqualified.

The difference between the proportion of jobs in which the NHW auditor does better and the proportion in which the minority auditor does better is generally taken as an indicator of discrimination. More precisely, it is an indicator of the presence of discrimination, not the extent of discrimination. If all employers were very discriminatory but both auditors were always totally unqualified for the jobs, there would be no jobs in which any auditor was hired and the other was not. So the size of the difference cannot be interpreted as a measure of the extent of discrimination. Often studies try to draw inferences by comparing the results for more skilled and less skilled jobs, or jobs sought by male auditors and jobs sought by female auditors, for instance, but such inferences have no basis.

**Table 26.1 Results of an employment audit
(Pager–Western–Bonikowski 2009)**

	NHW-Latino	NHW-Black
	% of employers	
1. Both get positive response	20.5%	14.7%
2. Neither gets positive response	65.5	71.3
3. NHW yes, minority no	10.6	16.4
4. NHW no, minority yes	4.7	0.6
Difference between 3 and 4	5.9%	15.8%

Pager et al. (2009) provide a good example of a traditional employment audit; it was performed by sociologists after economists had stopped doing such studies, and improves on previous studies in many small ways. In the part of the study I emphasize here, they randomly selected 170 employers looking for entry-level, low-skill workers in New York City, and sent teams of three auditors—one NHW, one Hispanic, and one black—to each of them.³ The auditors were carefully selected, carefully trained, and closely monitored young men alike in height, weight, education, and interpersonal skills. Table 26.1 (reproducing their table A-2) summarizes the results: The NHW-black difference was statistically significant, but the NHW-Hispanic difference was not.

The evidence for disparate treatment discrimination against blacks among these employers is fairly strong, but we cannot identify which employers are discriminators. Whether anyone is harmed is also unknown. The African American auditors were not harmed because they were not really looking for these jobs. Pager et al. (2009) point out that a black job seeker who randomly responded to newspaper and Craigslist advertisements would expect to take twice as long to find a job as an NHW job seeker following the same strategy, but black job seekers would not be wise to follow this strategy in New York City. In fact, only 14.1 percent of actual unemployed black job seekers in 2016 answered advertisements; considerably more (24.3 percent) contacted friends and relatives (US Bureau of Labor Statistics 2017: Table 33).

Among economists, traditional audits have fallen out of favor because of their small sample size (the 15.8 percent difference in the Pager et al. study represents about 27 employers), and the difficulties of making sure that auditors have nearly identical hard-to-observe characteristics that might matter to employers (such as personality), and that they do not try to tailor results to their political feelings, either consciously or unconsciously. To solve these problems, Bertrand and Mullainathan (2004) pioneered résumé audits: Instead of training auditors, they manufactured resumé to send to employers, and so they eliminated any problems with human auditors. To proxy for race, they manufactured white-sounding and black-sounding names from Massachusetts birth certificates, randomly attached them to resumé, matched similar resumé and

³ In another part of the study, some auditors claimed that they had criminal records. Whites with criminal records fared better than blacks without them.

sent them to employers. In 2001 and 2002, they sent over 5000 resumé's responding to more than 1300 employment ads in Boston and Chicago. Their measure of success was whether an employer called back for an interview. Over 80 percent of resumé's got no response. White names got callbacks 9.65 percent of the time and black names 6.45 percent, which is a highly significant difference. At the extremes, 15.9 percent of employers called "Brad," but only 2.2 percent called "Aisha."

Once again, we have evidence of disparate treatment discrimination. But whether this employer behavior harms anyone remains at issue. Moreover, since all of the "African Americans" in the sample have distinctively black names, they are not representative of the African American population. In fact, the discrimination that Bertrand and Mullainathan found may be, at least in part, discrimination by background rather than by race. Fryer and Levitt (2004) find that children with distinctively black names come from poorer backgrounds than black children without such names, so it is possible that some employers use the name as information about background, not race.

26.3.2.2 *Disparate Impact Results*

The more common exercise that economists have engaged in to try to find labor market discrimination is to compare the determinants of NHW and minority labor market outcomes (usually wages) using regression.⁴ If, on average, minorities with the same productively relevant characteristics as NHWs earn lower wages or work less, then they are worse off, demonstrating disparate impact. But this exercise can tell little about the process by which the disparate impact was created, and whether that process looked like disparate treatment discrimination.

Ideally, then, disparate impact and disparate treatment analyses could complement each other: audits could say what was happening on the ground and regressions could say whether it made a difference to well-being (or at least earnings). In practice, the process does not work out so neatly. Audits generally look at employment, and regressions generally look at wages. People who do not have jobs do not have wages that can be observed directly.

Generally, a disparate impact study begins with some data set of individuals that will include their wages (or other labor market outcome), and a vector of some of their characteristics, including race. Then the researcher will fit an equation like

$$\ln w_i = \beta X_i + \sum_j \gamma_j R_{ij} Y_i + \varepsilon_i$$

where i denotes individuals, j denotes races, X_i is a vector of individual i 's non-racial characteristics including a constant, R_{ij} is a dummy variable equal to 1 if and only if individual i is a member of race j , Y_i is a sub-vector of X_i including a constant, β and γ_j are vectors of coefficients to be estimated, and ε_i is an error term. The sum over j excludes one race.

⁴ It would be good to use job satisfaction as a dependent variable, but studies that do this are rare.

In the simplest version of this equation, Y_i is only a constant, and so γ_j becomes the premium (if positive) or penalty (if negative) that workers of race j enjoy or suffer relative to the excluded race. We call this the different-intercept method. The other polar specification is also commonly used, where $X_i = Y_i$ and in essence a separate regression is fit for each race, and then comparisons are made at the mean value of X_i for one race or another. The resulting analysis is known as a Blinder-Oaxaca decomposition.

One serious difficulty with these studies is that if the supports of the characteristics for different races have little or no overlap—if very few of one race have characteristics like people in another race—then conclusions about discrimination can be highly sensitive to functional form. With two races (A and B) and one characteristic, for instance, it is easy to construct an example where with the same data, with one specification one can conclude that race A is discriminated against; with another specification, that race B is discriminated against with another specification; and with a third specification, that there is no discrimination.

Another major issue with wage studies is how people who are not observed with wages should be treated: people in prison, disabled, long term unemployed, and so on. As we saw, at any moment fewer blacks have observable wages than NHWs or Hispanics, and the people who are not working generally have lower predicted wages than the people who have observed wages. Studies that ignore people without observable wages understate the extent of discrimination.

The other major issues with disparate impact regressions deal with the choice of independent variables: what does it mean to have “the same productively relevant characteristics”? Ideally, these are the same characteristics that would matter to an employer who could not observe or infer anything about extraneous traits. One obvious error is to leave out some characteristic that is relevant and correlated with race. Leaving out a relevant characteristic can make the estimated extent of discrimination either too big or too small, depending on how the left out variable is correlated with race. For instance, if we are trying to study discrimination against blacks relative to NHWs and leave out educational attainment, we are likely to bias our estimate up (because NHWs have higher educational attainment, which raises earnings), but if we are studying discrimination against Asians relative to NHWs we are likely to bias our estimate down (Asians have more education than NHWs). Since no data set has all relevant variables, objectors to any result in this literature can always cite some variable that should have been included but was not.

While leaving a variable out is an obvious problem, including a variable that should not be there can also bias estimates of discrimination. Suppose a regression in a different-intercept model includes some variable that does not affect productivity but is correlated with race, such as purchases of sunblock or hours watching ice hockey on TV. This extraneous variable will pick up some of the influence of race and bias down the estimate of discrimination. A Blinder-Oaxaca analysis—or simply an analysis where the extraneous variable is interacted with race—can avoid this problem.

Fryer et al. (2013) have an interesting strategy for avoiding both problems. In a sample of people who recently lost jobs, they use previous wage as a market-based estimate of

productivity. In the raw data, blacks are hired at jobs with wages 40 log points less than whites, but controlling for previous wage still leaves a gap of 17 log points. Other controls on such relevant items as bargaining strategy still leave a 16 log point gap. The data are from the Great Recession and so are fairly recent, and a wide array of people were unemployed. (They are also from New Jersey, a Northern, blue state.)

The key untestable assumption is that previous wage equals productivity, or at least that it differs equally from productivity for both blacks and whites on average. If black wages were below productivity in the previous job but white wages were not, then the estimates understate discrimination. On the other hand, if blacks benefited from affirmative action or tastes for diversity on the previous job on average, the estimates would overstate discrimination.

More traditional approaches try to collect appropriate sets of independent variables. For employment rates, Fairlie and Kletzer (1998) look at displacement rates (leaving jobs) and re-employment rates. Black workers have higher rates of displacement (about 30 percent higher) and lower rates of re-employment; so they are more likely to be unemployed for both reasons. Traditionally observable characteristics account for part of these raw gaps, but a large fraction is unexplained.

For wages, many studies with traditional variables were done in the twentieth century, but not many in the twenty-first century. Altonji and Blank (1999), for instance, find unexplained gaps for black men relative to NHW men of 12–15 percent (occasionally up to 20 percent), somewhat smaller gaps for black women, and gaps of around 11 percent for Hispanics overall relative to NHWs. (These equations do not control for occupation, since labor market discrimination may influence what occupations people enter.)

Since the 1980s, attention has focused on the test scores from the Armed Forces Qualifying Test (AFQT). These test scores are available for several thousand young people in the National Longitudinal Study of Youth (NLSY), who took this test around 1980. Members of the NLSY were born between 1957 and 1964. The AFQT is a test of cognitive skill that the armed forces developed to place people in the right military job, which seems to work well for military purposes and is probably racially unbiased. The AFQT score does not reflect intelligence; rather, it is a measure of the skills or achievement or what you learned in high school. Basically, your score rises as you get more education.⁵ In the NLSY, NHWs get higher AFQT scores than blacks or Hispanics. In fact, the distributions of AFQT scores of NHWs and minorities do not overlap much—the condition that we saw could lead conflicting interpretations depending on what specification was used.

In many wage equations, including AFQT as an explanatory variable removes a large part of the wage gap between black and white men, and eliminates all or most of it for black women and Hispanics. Neal and Johnson (1996) is the most famous paper that gets results like this. Notice that this result does not say that race no longer matters. Many papers show that race matters in determining AFQT in late adolescence, holding

⁵ Unfortunately, Herrnstein and Murray (1994) call the AFQT score “intelligence.” It is not.

family background constant.⁶ Neal and Johnson merely argue that most of the damage is done by late adolescence. The controversy about AFQT is about *when* race matters, not *whether* race matters.

Many papers have challenged Neal and Johnson's results. Some have added variables that Neal and Johnson omitted (Goldsmith et al. 1998; Lang and Manove 2011; Black et al. 2013); some have looked at the different parts of AFQT separately (Rodgers and Spriggs 1996); some have tried to correct for the omission of individuals without observed wages (Johnson et al. 2000); and some have looked at the NLSY cohort later in life (Tomaskovic-Devey et al. 2005). The conclusion seems to be that AFQT measures something that explains a significant part of the wage gap, but not all of it. Better education would cause very large wage gains for minorities, almost certainly greater than any kind of labor market policy could achieve, but there is no sharp dichotomy between activities before entering the labor market and activities after entering the labor market.

This body of work indicates that the wages blacks and Hispanics are offered are less than those that NHWs are offered by more than usually can be explained on non-racial grounds (although the literature is not unanimous, especially for Hispanics), and that employment rates are generally lower. The audit findings suggest that employer discrimination may explain some of these disparities. Nothing suggests that whites face discrimination that seriously impinges on their deliberative rights, although they may face instances of discrimination.

26.3.2.3 *What Do Civil Rights Laws Do?*

The policy implication from the research on discrimination is that civil rights laws in the labor market continue to have a role to play for blacks and Latinos. The problem that they are supposed to be addressing is real—perhaps not so important as educational disparities, but real nonetheless. However, results about the major racial anti-discrimination laws and affirmative action programs are in fact scarce, and high quality results are even more so.

The reason for this is practical: evaluation requires variation, and in the relevant period of abundant microdata, little variation in the relevant laws has occurred. The Civil Rights Act (containing Title VII) was signed in 1964, and although there have been amendments and significant court decisions since then, none has changed the landscape enough for economists to think the effects would be noticeable. (State civil rights laws vary, but the variation is likely to be endogenous.)

So the 1960s present the best opportunity to evaluate racial anti-discrimination and affirmative action policies, and most of the high quality results come from that decade. During the 1960s, the gap between black and white male wages narrowed—and after 1970, the gap stayed roughly the same for the rest of the twentieth century, once selection

⁶ If in every generation black children grew up with lower skills than white, conditional on family income, and blacks earned less than white, conditional on skills, then the black-white wage gap would always be growing over time. History shows that the wage gap has not always grown. If one wants to define “post-racial” as meaning “not everything gets worse for blacks all the time,” then the United States is post-racial. But most people probably do not define “post-racial” like this.

into employment is accounted for (Chandra 2003). (The pattern of women's wages is harder to discern because of the large changes in labor force participation.) The rise in black men's wages was concentrated in the South, and coincided with the entry of large numbers of black men into Southern manufacturing, which previously had been essentially all white. Careful analyses of these patterns supports the hypothesis that federal civil rights laws and affirmative action programs were responsible for much of this narrowing of the wage gap (Heckman and Paynter 1989). The approximate constancy of the gap over the next three decades is also consistent with the approximate constancy of these policies.

The reduction in the black-white wage gap in the 1960s also seems to have been a potential Pareto improvement (the winners could have more than compensated the losers, even if they did not), at least as conventionally measured (that is, without considering vicarious losses to racists). Southern manufacturers appeared to have wanted to hire blacks, but were afraid of a backlash from white governments and thugs. Federal pressure gave them cover to do what they wanted to do, including operating integrated cafeterias and restrooms. Manufacturers and black workers gained, while owners of farmland and associated capital lost, but since black workers were moving to a higher return activity, the aggregate losses were less than the aggregate gains.

Although racial civil rights laws appear to have worked well in the 1960s, there is no *a priori* assurance that anti-discrimination laws in general either help their intended beneficiaries or cause potential Pareto improvements. Employers can evade them, which can be socially costly in unforeseen way, and they can also be extremely hard to enforce, even if employers do not evade them. Edelman (2016) provides a detailed study of how anti-discrimination laws work in large corporations and argues that corporations have coopted judges, resulting in an emphasis on policies and procedures that are mostly symbolic and fail to produce results.

We can also see some of these effects in non-racial anti-discrimination laws implemented after the 1960s. The Americans with Disabilities Act (ADA) of 1990 forbids discrimination based on disability and requires employers of certain sizes to make reasonable accommodations. Although a few papers disagree, the weight of the evidence now is that the ADA reduced employment of disabled workers, especially with large employers (Acemoglu and Angrist 2001). The reasonable accommodation requirement seems to have acted like a tax on the employment of disabled workers.

States have enacted many laws governing drug testing by employers, both for and against. The more that states require drug testing, the more black men are employed; the more that states prohibit drug testing, the fewer black men are employed (Wozniak 2014). Apparently, drug testing gives black men who do not use drugs a way of proving to potential employers that they do not, and this is valuable for them.

Similarly, emerging evidence (Doleac and Hansen 2016) shows that ban-the-box laws also reduce the employment of young black and Hispanic men. These are laws that forbid employers from asking whether a job applicant has a criminal record until late in the application process. The idea is to give people with criminal records a better shot at finding employment. Apparently, some employers suspect every young minority man of

being a criminal, and ban-the-box laws keep minority men without criminal records (the majority) from being able to prove that they are more qualified than employers think.

26.4 POLICIES FOR THE FUTURE

Both economists and philosophers should think about alternatives to civil rights law. The long stagnation of the black-white wage gap that started around 1970 suggests that these laws by themselves will never assure deliberative rights (except possibly in geological time).

Affirmative action is the obvious alternative: mandating that employers hire certain numbers of black and Latino workers. Relative to civil rights laws, affirmative action is much less intrusive: employers do not have to justify their actions to anyone as long as they hire and retain enough minorities. They can also choose unhindered the minority workers whom they think are most productive and most compatible with their firm, not those who are most litigious. But affirmative action is controversial, and a full discussion is beyond the scope of this paper.

Another approach is to think of current civil rights laws, not only as reducing the amount of information that employers can legally use, but as increasing the amount of job-relevant information they have. Consider an employer who discriminates for statistical reasons. In the original situation, absent civil rights laws, she cares about some job-relevant measure of qualification q , but can observe only race, which is correlated with q . Then she will base all her decisions on race. Now suppose that a sort of “radar” gun is invented that, when pointed at an individual, it reveals q . Given q , employers have no interest in race, and so an employer with such a radar gun has no incentive to discriminate on the basis of race. A government interested in eliminating discrimination can either adopt an anti-discrimination law that it tries to enforce, or distribute radar guns to employers. The anti-discrimination law will make firms less productive, while the radar gun will make them more productive: Employers will try to evade the anti-discrimination law, but they will embrace their radar guns.

This is, of course, the strongest possible case for radar guns. Radar guns will not affect taste-based discrimination. Nor will they assuage employer fears about poor interactions with fellow employees and customers (although these might be mitigated if all firms got radar guns at the same time). But radar guns can exist side-by-side with anti-discrimination laws, and reduce incentives to evade those laws.

The most serious objection is that radar guns are science fiction. But improved techniques, such as tests, to enable employers to learn about qualifications more directly are not science fiction. The federal government should invest in research to improve the techniques that employers use to judge the qualifications of applicants and employees. This research is a public good: the federal government, as the largest employer in the nation, can use these techniques to improve its own work force, and many kinds of research will be useful for other employers. Use of effective screening tools by a single

employer can also have spillover benefits for other employers, who can infer something about the qualifications of applicants by observing whether the firm with effective screening tool made an offer to them. (Perhaps, then, deliberative freedom includes the freedom to make one's qualifications known.)

26.5 CONCLUSION

What do civil rights laws do? They assign certain rights to certain people. In societies where some groups have suffered pervasive discrimination, there are good arguments to assign deliberative rights that are often held by employers to jobseekers and workers who are members of groups experiencing pervasive discrimination. These arguments do not apply particularly well to jobseekers and workers from other groups. In modern America, economists' disparate impact studies indicate fairly convincingly that African Americans and Latinos meet the conditions for coverage by civil rights laws, but NHWs do not.

These arguments for assigning deliberative rights to members of groups who are victims of pervasive discrimination apply not only to the supply of labor, but to many other areas as well. Some of these areas, like public accommodations and education, are covered by civil rights laws today, while other areas, like adoption and the demand side of employment, are not.

Like all rights assignments, civil rights laws sometimes require people to do things that they do not want to do, and to refrain from doing things that they want to do. For that reason, these laws require enforcement, and because the rights are defined by a counterfactual, enforcement is imperfect. The people whom these laws force to do things they would not do otherwise include not only racists who bear animosity to the protected groups, but also people with completely standard utility functions—strategic discriminators—who want to discriminate because it would help them achieve standard goals like profit maximization, crime reduction, or promotion.

Because enforcement is imperfect, and because strategic discriminators are present, civil rights laws in one domain are complements to civil rights laws in other domains. Because of increasing returns to scale, efficient reduction of discrimination in any one domain in general requires reduction of discrimination in all domains. Because discrimination in some domains is unlikely ever to be circumscribed by civil rights laws, employment-related civil rights laws might be needed even if no employers were taste-based discriminators. Having a few civil rights laws does not guarantee that anyone has her civil rights in any domain.

Since 1970 or so, moreover, there is no hard data to suggest that civil rights laws in employment have had much success in improving the deliberative freedoms of African Americans and Latinos, and some qualitative legal analysts have suggested they are ineffective. Something more is needed. One possibility is affirmative action, which is simple and much easier to administer than civil rights laws, but controversial and beyond

the scope of this paper. I have suggested some alternatives that use more information rather than less.

What is most distressing is that neither economists nor philosophers are spending much effort on finding better ways to assure deliberative freedoms to African Americans and Latinos. The data are pretty clear that the legislation of the 1960s was not sufficient. Keeping our thoughts in the 1960s is not sufficient either.

ACKNOWLEDGMENTS

I am grateful for helpful comments from Philip S. Kitcher, Rajiv Sethi, and Mark D. White. But they are not responsible for the errors.

REFERENCES

- Acemoglu, Daron, and Joshua Angrist. 2001. "Consequences of Employment Protection: The Case of the Americans with Disabilities Act." *Journal of Political Economy* 109: 915–57.
- Altonji, Joseph, and Rebecca Blank. 1999. "Race and Gender in the Labor Market." In Orley C. Ashenfelter and David Card (eds.), *Handbook of Labor Economics*, vol. 3 (Amsterdam: North-Holland), pp. 3143–259.
- Autor, David, and Philip S. Kitcher. 2018. "As You Like It: Work, Life, and Satisfaction." In Subramanian Rangan (ed.), *Capitalism Beyond Mutuality? Perspectives Integrating Philosophy and Social Science* (Oxford: Oxford University Press), pp. 139–60.
- Becker, Gary. 1957. *The Economics of Discrimination*. Chicago, IL: University of Chicago Press.
- Bertrand, Marianne, and Sendhil Mullainathan. 2004. "Are Emily and Greg More Employable than Lakisha and Jamal? A Field Experiment on Labor Market Discrimination." *American Economic Review* 94: 991–1013.
- Black, Dan A., Natalia Kolesnikova, Seth G. Sanders, and Lowell J. Taylor. 2013. "The Role of Location in Evaluating Racial Wage Disparity." *IZA Journal of Labor Economics* 2(2). Available at <http://www.izajole.com/content/2/1/2>.
- Chandra, Amitabh. 2003. "Is the Convergence of the Racial Wage Gap Illusory?" National Bureau of Economic Research Working Paper 9476. Available at <http://www.nber.org/papers/w9476>.
- Doleac, Jennifer, and Benjamin Hansen. 2016. "Does 'Ban-the-Box' Help or Hurt Low-Skilled Workers? Statistical Discrimination and Employment Outcomes When Criminal Records Are Hidden." National Bureau of Economic Research Working Paper 22469. Available at <http://www.nber.org/papers/w22469>.
- Durlauf, Steven N. 2006. "Assessing Racial Profiling." *Economic Journal* 116: F402–F426.
- Edelman, Lauren B. 2016. *Working Law: Courts, Corporations, and Symbolic Civil Rights*. Chicago, IL: University of Chicago Press.
- Fairlie, Robert W., and Lori G. Kletzer. 1998. "Jobs Lost, Jobs Regained: An Analysis of Black/White Differences in Job Displacement in the 1980s." *Industrial Relations* 37: 460–77.
- Fryer, Roland G., Jr, and Steven Levitt. 2004. "The Causes and Consequences of Distinctively Black Names." *Quarterly Journal of Economics* 119: 767–806.
- Fryer, Roland G., Jr, Devah Pager, and Jorg L. Spenkuch. 2013. "Racial Disparities in Job Finding and Offered Wages." *Journal of Law and Economics* 56: 633–89.

- Guardian, The. 2017. "The Counted: People Killed by Police in the US." Available at <https://www.theguardian.com/us-news/ng-interactive/2015/jun/01/the-counted-police-killings-us-database>.
- Goldsmith Arthur H, William Darity, Jr, and Jonathan R. Veum. 1998. "Race, Cognitive Skills, Psychological Capital, and Wages." *Review of Black Political Economy* 26 (2): 9–21.
- Heckman, James. 1998. "Detecting Discrimination." *Journal of Economic Perspectives* 12: 101–16.
- Heckman, James, and Brook S. Paynter. 1989. "Determining the Impact of Federal Antidiscrimination Policy on the Economic Status of Blacks: A Study of South Carolina." *American Economic Review* 79: 219–77.
- Herrnstein, Richard J., and Charles Murray. 1994. *The Bell Curve: Intelligence and Class Structure in America*. New York: Free Press.
- Johnson, William R., Yuichi Kitamura, and Derek Neal. 2000. "Evaluating a Simple Method for Estimating Black-White Gaps in Median Wages." *American Economic Review* 90: 339–43.
- Kitcher, Philip S. 1999. "Race, Ethnicity, Biology, Culture." In Leonard Harris (ed.), *Racism* (New York: Humanity Books), pp. 87–117.
- Lang, Kevin, and Michael Manove. 2011. "Education and Labor Market Discrimination." *American Economic Review* 101: 1467–96.
- Leeth, John D., and John Ruser. 2006. "Safety Segregation: The Importance of Gender, Race, and Ethnicity on Workplace Risk." *Journal of Income Inequality* 4: 123–52.
- Lippert-Rasmussen, Kasper. 2006. "The Badness of Discrimination." *Ethical Theory and Moral Practice* 9: 167–85.
- Lippert-Rasmussen, Kasper. 2013. "Discrimination." In Hugh LaFollette (ed.), *International Encyclopedia of Ethics* (Oxford: Wiley-Blackwell), pp. 1405–15.
- Moreau, Sophia R. 2010. "What is Discrimination?" *Philosophy and Public Affairs* 38: 143–79.
- Mukerjee, Swati. 2014. "Job Satisfaction in the United States: Are Blacks Still More Satisfied?" *Review of Black Political Economy* 41 (1): 61–81.
- Myrdal, Gunnar. 1944. *An American Dilemma, Volume I: The Negro Problem and Modern Democracy*. New Brunswick, NJ: Transaction Publishers (1996 edition).
- Neal, Derek A., and William R. Johnson. 1996. "The Role of Pre-labor Market Factors in Black-White Wage Differences." *Journal of Political Economy* 104: 868–95.
- O'Flaherty, Brendan. 2015. *The Economics of Race in the United States*. Cambridge, MA: Harvard University Press.
- Pager, Devah, Bruce Western, and Bart Bonikowski. 2009. "Discrimination in a Low-Wage Labor Market: A Field Experiment." *American Sociological Review* 74: 777–99.
- Rawls, John. 1971. *A Theory of Justice*. Cambridge, MA: Harvard University Press.
- Risse, Mathias, and Richard Zeckhauser. 2004. "Racial Profiling." *Philosophy and Public Affairs* 32: 131–70.
- Rodgers, William, III, and William E. Spriggs. 1996. "What Does AFQT Really Measure? Race, Wages, Schooling, and the AFQT Score." *Review of Black Political Economy* 24 (4): 13–46.
- Shin, Patrick. 2013. "Civil Rights." In Hugh LaFollette (ed.), *International Encyclopedia of Ethics* (Oxford: Wiley-Blackwell), pp. 794–803.
- Steele, Claude M. 2010. *Whistling Vivaldi: And Other Clues to How Stereotypes Affect Us*. New York: W.W. Norton & Company.
- Taylor, Paul C. 2013. *Race: A Philosophical Introduction*. 2nd edn. Cambridge: Polity Press.
- Tomaskovic-Devey, Donald, Melvin Thomas, and Kecia Johnson. 2005. "Race and the Accumulation of Human Capital across the Career: A Theoretical Model and Fixed Effects Application." *American Journal of Sociology* 111: 58–89.

- US Bureau of the Census. 2017. CPS Creator. Available at <https://www.census.gov/cps/data/cpstablecreator.html>.
- US Bureau of the Census. 2018. "Historical Income Tables." Available at <https://www.census.gov/data/tables/time-series/demo/income-poverty/historical-income-households.html>.
- US Bureau of Labor Statistics. 2017. "Labor Force Statistics from the Current Population Survey." Available at <https://data.bls.gov/cgi-bin/surveymost?ln>.
- Viscusi, W. Kip. 2003. "Racial Differences in the Labor Market Values of a Statistical Life." *Journal of Risk and Uncertainty* 27: 239–56.
- Wozniak, Abigail. 2014. "Discrimination and the Effects of Drug Testing on Black Employment." National Bureau of Economic Research Working Paper 20095. Available at <https://www.nber.org/papers/w20095>.

PART III

CONCLUSION

CHAPTER 27

LESSONS FROM ECONOMICS

JOHN BROOME

27.1 INTRODUCTION

ECONOMICS and the methods of economics offer several lessons that moral philosophers could beneficially learn. I shall mention some of them, using two significant examples of topics where they can be useful. Indeed, on these particular topics, ignoring the lessons from economics has been seriously damaging; I shall show that it has led to bad mistakes. So my conclusion is not merely that lessons from economics could be beneficial. They are actually essential for dealing properly with some topics within moral philosophy.

Much of moral philosophy is concerned with the personal relations and interactions of a few people. Economics has little to contribute there. But many moral philosophers these days deal with large-scale subjects that have quantitative aspects. For example, they deal with the ethics of public health, of social inequality, and of population. These subjects concern economists too. A branch of the discipline of economics known as “welfare economics” is concerned with making ethical judgments about them and other subjects.

Public health, social inequality, and population each have a quantitative aspect and an ethical aspect. Economists are adept at analysing the quantitative aspects using mathematical methods. To be successful, their methods must be based on good ethical theory. But the academic labor cannot just be divided, with moral philosophers working on the ethical bases of economics, and economists applying philosophers’ conclusions to complex, large-scale problems. The ethics and the economics are too closely intertwined to be teased apart.

Each discipline must work with the other. Economists typically have their own moral philosophy. For example, they typically believe that value derives only from people’s

preferences. Partly as a result, some of them have the surprising view that different people's well-beings cannot be compared together. Philosophy can help them clarify their thinking. The papers in this handbook describe some of the lessons that economists should learn from philosophy.

Conversely, moral philosophers must pay attention to the methods and conclusions of economics. That was the theme of my book *Ethics Out of Economics* (Broome 1999); I return to it now with new examples. Unfortunately, moral philosophers have been slow to learn from economists. This is partly because some of them are unwilling to employ or even read mathematical formulae, which are the staple of economics. No one need engage in a discipline that does not interest her, but a philosopher who does not engage in mathematics should avoid certain quantitative subjects within moral philosophy. When economists employ mathematics, they usually do so because it is needed.

To illustrate the lessons moral philosophy can learn from economics, I have chosen two particular topics as examples: the ethics of equality and the ethics of population. Sections 27.3–27.6 describe how the methods of economics can contribute to our understanding of the value of equality. Section 27.7 illustrates how they can help with population ethics and some connected topics.

Those are examples of beneficial lessons that philosophers can learn from economics. But this chapter starts in Section 27.2 with a bad lesson from economics that many moral philosophers have already thoughtlessly learnt. It is the use of the word “utility” to denote a person's good or well-being. This usage was invented by economists. They would have done better not to have invented it, and moral philosophers would do better not to copy them. For some reason, many philosophers seem more attracted by this bad habit of economists than by their good ones.

27.2 DO NOT USE “UTILITY” TO DENOTE A PERSON'S GOOD

It is perfectly unnecessary to use the word “utility” to denote a person's good or well-being. We already have the words “good” and “well-being.” “Benefit” is another useful word with the meaning of “add to good.” For explicit comparisons we have “better for.” We do not need a further artificial term. Moreover, using “utility” with this meaning leads to mistakes. It has been doing damage in economics since the middle of the twentieth century, and more recently in philosophy too. I first complained about this usage in Broome (1991a), at a time when few philosophers had adopted it. But by now many have. It should be eschewed.

In economics, “utility” is ambiguous. First, it may refer to something that plays a role in economists' ethical judgments or their judgments of value. Much of economic theory is concerned with making valuations of states of affairs and policies, to judge how good or bad they are. Often these valuations are based on a “social welfare function,” which

could alternatively be called a “value function.” Typically, a social welfare function is written as a function of people’s utilities: $W(u_1, u_2, \dots, u_n)$, where u_i stands for the i th person’s utility. If the function is specifically the sum of utilities $W = u_1 + u_2 + \dots + u_n$, it is referred to as a “utilitarian” function. Utilitarianism is the view that the value of a state of affairs is the sum of individuals’ well-beings, so this tells us that utility is understood to be well-being. Even when the function is not additive, the role of a person’s utility within a social welfare function is to denote her well-being or what is good for her. That is one meaning of “utility.”

Mistakes occur because the same word is also used by economists with a quite different meaning. “Utility” is defined in economic theory as the value of a function that represents a person’s preferences. This function *represents* preferences in the sense that one thing is defined to have greater utility than another if and only if the person prefers it. This means that a person’s preferences maximize utility, by the definition of utility. Utility defined this way is everywhere in economic theory. The definition appears in the canonical works such as Debreu (1959) and Deaton and Muellbauer (1980). It plays an important role in consumer theory, general equilibrium theory, game theory, and all those other parts of economics where a person’s behavior is supposed to be given by her preferences. In none of these places need there be any implication that utility is a person’s good. “Utility” in these contexts is a totally different, technical term.

However, economists tend to assume that one thing is better for a person than another if and only if she prefers it. This may be because they hold the theory that a person’s good consists in the satisfaction of her preferences, or because they assume that a person determines her preferences on the basis of what is better for her. But whatever the reason, their view is no doubt encouraged by their inclination to use “utility” ambiguously for good and also for a representation of preferences.

That is damaging insofar as economists are wrong to assume that one thing is better for a person than another if and only if she prefers it. Much more serious damage arises when they come to use the more restrictive technical definition of utility that is found in expected utility theory. There, “utility” is defined to mean, not the value of any function that represents a person’s preferences, but specifically the value of a function that represents a person’s preferences in an *expectational* way. Let a “prospect” be the set of possible outcomes that might result from some event, each associated with its probability of occurring. In expected utility theory, the utility of a prospect is defined to be the mathematical expectation of the utility of its possible outcomes. Because, by definition, a person’s preferences maximize her utility, this further definition means that her preferences over prospects maximize her expected utility. Again, utility defined this way plays an important role in economists’ theory of people’s behavior—this time, their behavior in the face of uncertainty. Again, there need be no implication that a person’s utility is her good.

Utility defined this way, which I shall call *expectational utility*, has the property of being *cardinal*. Many different utility functions can represent a person’s preferences. But all the functions that represent her preferences expectationally are related to each other in a particular way: they are all affine transforms of each other.

A utility function $u'()$ is defined to be an *affine transform* of another $u()$ if and only if

$$u'(x) = au(x) + b \text{ for all } x, \text{ where } a \text{ and } b \text{ are constants and } a \text{ is positive.}$$

Utility functions that are affine transforms of each other may differ in what state of affairs they assign zero utility to, and they may differ in the size of the unit of utility, but in other ways they do not differ. In particular, they do not differ in the ratio of utility differences. If the difference in utility between A and B is twice the difference between C and D according to one utility function, that is also so according to every utility function that is an affine transform of it. This means that ratios of differences are significant in an expectational utility function, whereas the zero and unit are not. This is what it means to say that expectational utility is cardinal.

Many economists assume that a person's utility defined this way is her good. More exactly, they assume it measures her good cardinally, by which I mean it is an affine function of her good. That is: $u(w) = aw + b$ where w is the person's good or well-being, u her utility, a and b are constants and a is positive. Why do they assume this? On the face of it, it is very implausible. It implies that the goodness of a prospect for a person is just the expectation of the good she may derive from the prospect's various possible results. This is to rule out the possibility that risk to her good is bad for a person.

For example, it implies that getting one unit of benefit for sure is equally as good for a person as a gamble that gives her two units or no units with equal probability. If, say, the goodness of a day in a person's life is given by the number of hours she spends fiddling with her phone, or the number of hours she spends thinking through deep problems in philosophy, then an hour for sure engaged in one of these activities is no better than a risky gamble that gives her two hours or no hours with equal probability. This is not an intuitively attractive assumption. Quite plausibly, the risk-free option may be better.

A way of putting it is that risk to good is "neutral" according to the view that utility measures good cardinally. There is a theoretical defense of this view, which I shall describe at the end of Section 27.5, but this defense was not available to economists when they first adopted this view. The most likely explanation of the economists' assumption is they were confused by their own terminology. By the definition of utility as a cardinal representation of a person's preferences, preferences are neutral about risk to utility. Since economists also use "utility" for a person's good, they assume risk to good is neutral.

There is no need for this assumption. A person's utility need not be an affine transform of her good. Instead, it could be a strictly concave function of her good (which means that the graph of utility against good curves downwards as it slopes upwards). Then risk to good will be bad, not neutral.

The assumption that risk to good is neutral is an old one. Daniel Bernoulli (1738) seems to have made it in his account of the St Petersburg paradox, which may fairly be considered the beginning of expected utility theory. For this reason, I call this assumption "Bernoulli's Hypothesis." Bernoulli was not confused by terms. He wrote in Latin, and what he actually assumed is that risk to *emolumentum* is neutral. *Emolumentum* is badly translated as "utility" in the 1954 *Econometrica* version of the paper. "Good," "benefit," or "well-being" would be better translations. Bernoulli's achievement was to

show that expected utility theory does not have to assume that risk to money is neutral. A hundred dollars for sure may be better than a gamble at equal odds on \$0 and \$200; this is consistent with the theory. It is an easy next step to realize that the theory also does not have to assume that risk to well-being or *emolumentum* is neutral. Bernoulli happened not to make that step.

Nor did anyone for the next two centuries, it seems. Bernoulli's Hypothesis persisted at the foundation of expected utility theory for all that time. Exactly because it is intuitively unattractive, expected utility theory fell into disrepute among economists in the first half of the twentieth century (see for instance Tintner 1942).

It was spectacularly revived by John von Neumann and Oskar Morgenstern's (1944) formulation of axiomatic expected utility theory. (Frank Ramsey's earlier formulation in 1931 received little attention.) What von Neumann and Morgenstern showed is that, provided a person's preferences conform to some axioms, she can be treated as an expected-utility maximizer. That is to say, a utility function can be defined for her that represents her preferences in an expectational way. Nothing in von Neumann and Morgenstern's work suggests that this utility measures the person's well-being. Their work gives no support to Bernoulli's Hypothesis, though they themselves did not make this clear.

Some clear-minded economists saw it immediately. For example, Kenneth Arrow says:

[von Neumann and Morgenstern's] theorem does not, as far as I can see, give any special ethical significance to the particular utility scale found. For instead of using the utility scale found by von Neumann and Morgenstern, we could use the square of that scale; then behavior is described by saying that the individual seeks to maximize the expected value of the square root of his utility. This is not to deny the usefulness of the von Neumann-Morgenstern theorem; what it does say is that among the many different ways of assigning a utility indicator to the preferences among alternative probability distributions, there is one method (more precisely, a whole set of methods which are linear transforms of each other) which has the property of stating the laws of rational behavior in a particularly convenient way. This is a very useful matter from the point of view of developing the descriptive economic theory of behavior in the presence of random events, but it has nothing to do with welfare considerations, particularly if we are interested primarily in making a social choice among alternative policies in which no random elements enter. To say otherwise would be to assert that the distribution of the social income is to be governed by the tastes of individuals for gambling. (1951: 10)

Other economists drew the opposite conclusion. They concluded that expectational utility measures well-being or good. This means they implicitly adopted Bernoulli's Hypothesis, however implausible it may seem. John Harsanyi is a leading example. This remark of his illustrates his use of "utility":

If I want to compare the utility that I would derive from a new car with the utility that a friend would derive from a new sailboat, then I must ask myself what utility I would derive from a sailboat if I had taken up sailing for a regular hobby as my friend has done, and if I could suddenly acquire my friend's expert sailing skill, and so forth. (1977: 59)

Harsanyi evidently means to ask himself how much benefit he would derive from a new sailboat if he had taken up sailing. He is using “utility” for benefit or good. This earlier passage gives a fuller description of his thinking:

To be sure, the vNM utility function of any given individual is estimated from his choice behavior under risk and uncertainty. But this does not mean that his vNM utility function is *merely* an indication of his attitude towards risk taking. Rather, as its name shows, it is a utility function, and more specifically, it is what economists call a cardinal utility function. This means that the primary task of a vNM utility function is *not* to express a given individual’s attitude toward risk taking; rather it is to indicate how much utility, i.e. how much subjective *importance*, he assigns to various goals... Consequently, vNM utility functions have a completely legitimate place in ethics because they express the subjective importance people attach to their various needs and interests. (1975: 600)

This argument is evidently founded on what Harsanyi takes to be the meaning of the terms “utility” and “cardinal utility.” He writes as though he is describing the ordinary meaning of “utility” in English, but he mistakes economists’ jargon for real English. In English “utility” means usefulness, but that is not the meaning Harsanyi describes. He does not even give us the meaning of “utility” that is officially defined in economics: the value of a function that represents preferences. Instead, he says that “utility” *means* subjective importance. Because he is evidently a subjectivist about a person’s well-being, he takes “utility” to refer to well-being.

He denies that the primary task of a vNM utility function is to express a person’s attitude towards risk. But that is indeed its primary function. It is a technical notion that is defined to have exactly this function within von Neumann and Morgenstern’s theory. We could understand it as referring to a particular sort of subjective importance for the special purpose of decision making under risk. But as Arrow pointed out 25 years earlier, nothing in the definition implies that “utility” refers to subjective importance in any way that is relevant to ethics. Harsanyi’s thinking seems to have been driven by a misunderstanding about the meaning of “utility.” In this he is typical of many economists.

I shall explain at the end of Section 27.5 that Harsanyi had available a much better argument against Arrow. Moreover, this argument derives from Harsanyi’s own work, as we shall see. However, Harsanyi himself chose to rely on this weak argument based on the meaning of “utility.”

Now to just one example of modern practice in moral philosophy. In “Why It Matters That Some Are Worse Off than Others,” Michael Otsuka and Alex Voorhoeve (2009: 173) explicitly take “utility” to refer to how well a person’s life is going, and they also assume that utility is expectational (2009: 172–3n3). So these authors assume that expectational utility measures well-being. That is to say, they commit themselves unhesitatingly to Bernoulli’s Hypothesis, despite its implausibility. This seems unwise, given that their aim in this paper is to refute prioritarianism. We shall see at the end of Section 27.5 that the question of how to measure well-being—in particular whether to use expectational utility as a measure—is a central issue in judging whether prioritarianism is credible.

27.3 THE DEFINITION OF PRIORITARIANISM

After the bad lesson, I come to some good lessons that philosophers can learn from economics. I start with all the large number of particular truths that economists have discovered about topics that interest philosophers. These truths mostly lie within the theory of value. The theory of the value of equality is the best example. Economists have studied the value of equality for a long time. Philosophers came more recently to the subject, and can make use of discoveries economists made previously.

In 1991, Derek Parfit gave a Lindley Lecture at the University of Kansas entitled “Equality or Priority.” He introduced what he called “the priority view” and distinguished it from what he called “egalitarianism.” This lecture is described as the *locus classicus* of the priority view by Otsuka and Voorhoeve (2018), though these authors do recognize economists’ earlier work on the subject. In a footnote, Parfit (1991: 41n30) referred to some philosophers who had earlier supported the priority view, but not to any economist who had done so.

He appears not to have known that the priority view was sufficiently popular in economics that it could fairly be called the economists’ standard account of the value of equality. For example, it is explicitly stated in a standard textbook by Anthony Atkinson and Joseph Stiglitz (1980: 340). However, I believe the useful name “the priority view” was Parfit’s invention. It has now given way to the equally useful “prioritarianism.”

Much earlier, economists had identified a precise way of making the distinction between the priority view and others. It is best expressed as a condition on a betterness ordering among distributions of well-being. A distribution can be described by a vector of people’s well-beings $w = (w_1, w_2, \dots, w_n)$. In the literature on equality, it is nearly always assumed that people’s well-being has an interpersonally comparable cardinal measure, and I shall continue with that assumption. So I take each w_i to be a real number. I assume the population of people is constant at n . Betterness is assumed to constitute a *complete ordering* (a transitive and asymmetric relation) on the set of distributions. I am speaking of “overall” or “general” betterness: betterness from the point of view of society or the universe. We can specify various condition on this betterness ordering.

For example, we might expect it to be *increasing* in each person’s well-being. That is:

$(w_1, w_2, \dots, w_n)$ is better than $(w'_1, w'_2, \dots, w'_n)$ if
for some i , $w'_i > w_i$ and for all $j \neq i$, $w'_j = w_j$.

We might also expect it to be *impartial* between people, which means that:

$(w_1, w_2, \dots, w_n)$ is equally as good as $(w'_1, w'_2, \dots, w'_n)$ if
 $(w_1, w_2, \dots, w_n)$ is a permutation of $(w'_1, w'_2, \dots, w'_n)$.

The *Pigou-Dalton condition* is a satisfactory way of specifying the condition that equality is valuable. It says that a transfer of well-being from a better-off person to a less well-off

one makes the distribution better, provided the transfer is not enough to reverse the people's relative positions. That is:

$(w_1, w_2, \dots, w_n)$ is better than $(w'_1, w'_2, \dots, w'_n)$ if,
 for some i and some j , $w_i + w_j = w'_i + w'_j$, and $w'_i > w_i \geq w_j > w'_j$, and
 for all k such that $k \neq i$ and $k \neq j$, $w_k = w'_k$.

I also need to formulate a condition of *separability*. Let I be a subset of the population made up of the people $i, j, \dots, k$. Let $w_I = (w_i, w_j, \dots, w_k)$ be the distribution of well-being over these people, and let w_R be the distribution over the rest of the population. Given impartiality, the order of elements does not matter, so we may write the whole distribution as (w_I, w_R) . Compare the four distributions (w_I, w_R) , (w'_I, w_R) , (w_I, w'_R) , and (w'_I, w'_R) . The first pair differ only in the well-beings of the members of I ; the second pair also differ only in the well-beings of those same people. Furthermore, the difference between the first and second member of each pair is exactly the same in each case. The subset of people I is defined to be *separable* in the betterness ordering if and only if, for all distributions, (w_I, w_R) is better than (w'_I, w_R) if and only if (w_I, w'_R) is better than (w'_I, w'_R) . Separability means that the influence of the subset I on the ordering is independent of whatever the rest of the distribution may be. The betterness ordering is defined to be *strongly separable* if and only if every subset of the population is separable in the ordering.

I shall deal only with theories of betterness that imply an increasing and impartial betterness ordering. Of these, the ones that give value to equality also satisfy the Pigou-Dalton condition. These latter theories can be divided into two classes. Those that satisfy strong separability are prioritarian. Those that do not Parfit calls "egalitarian," but I shall call them "nonprioritarian."

I have presented this definition of prioritarianism in terms of the structure of the betterness ordering rather than in terms of a value function. This is because a value function does not always exist. A value function is a way of representing an underlying betterness relation in terms of real numbers. Given a betterness relation, a value function $v(w)$ *represents* it if and only if, for any distributions w and w' , $v(w) > v(w')$ if and only if w is better than w' . Not all betterness orderings can be represented by a value function. For instance, some orderings with lexical features cannot be. Also, an ordering that can be represented does not have a unique representation; it can be represented by many different value functions.

When a betterness ordering does have a value function, it is strongly separable if and only if it is *additively separable* (Debreu 1954). An additively separable ordering is one that can be represented by a value function having the form:

$$v(w) = v_1(w_1) + v_2(w_2) + \dots + v_n(w_n) \quad (1)$$

(It will also have other representations that do not take this form.) If the ordering is to be increasing and impartial and satisfy the Pigou-Dalton condition, each of the functions $v_i()$

must be the same increasing, strictly concave function. $v_i(w_i)$ can be thought of as the contribution of i 's well-being to general value $v(w)$. Equation (1) clearly shows how each person's well-being makes a contribution to general value that is independent of other people's well-beings. This is the characteristic of prioritarianism that Parfit stresses.

Nevertheless, we cannot define prioritarianism in terms of the additively separable form of the value function, because some betterness orderings that should be counted as prioritarian cannot be represented by a value function. One of these is the so-called *leximin* ordering, which holds an important place in the history of prioritarianism. It is defined by the criterion:

One distribution w is better than another w' if and only if:

- the person with the lowest well-being in w is better off than the person with the lowest well-being in w' , or
- the person with the lowest well-being in w is equally as well off as the person with the lowest well-being in w' , and the person with the second-lowest well-being in w is better off than the person with the second-lowest well-being in w' , or
- the person with the lowest well-being in w is equally as well off as the person with the lowest well-being in w' , and the person with the second-lowest well-being in w is equally as well off as the person with the second-lowest well-being in w' , and the person with the third-lowest well-being in w is better off than the person with the third-lowest well-being in w' , or
- ...and so on.

The leximin ordering satisfies the Pigou-Dalton condition and it is strongly separable. So it is prioritarian by the definition I gave. This is a desirable conclusion. The leximin ordering was introduced by Amartya Sen (1970: 138n) in discussing the "maximin" ordering, which he ascribed to John Rawls (later published in Rawls 1971). (The maximin ordering has the value function $\min(w_1, w_2, \dots, w_n)$. It does not satisfy the Pigou-Dalton condition.) In an appendix to the Lindley Lecture, Parfit (1991) argues that the leximin ordering, rather than the maximin one, represents Rawls's true view about equality. Parfit uses informal arguments to show leximin is prioritarian. Historically, the leximin theory is one of the more prominent prioritarian theories.

Yet it cannot be represented by a value function, whether additively separable or not. So we have to use the broad definition of prioritarianism that I gave initially, in terms of the structure of the betterness relation. This is an example of a useful general lesson moral philosophy can learn from economics. Philosophers often start their arguments from their intuitions about value. But betterness is generally more fundamental than value, and a better place to start. (This is not to deny that particular values often contribute to determining what the betterness ordering is.) Not only are many betterness orderings not represented by value functions, but intuitions about value may not have a clear meaning unless they are anchored in intuitions about betterness. We shall come to an example in Section 27.7.

27.4 A BAD ARGUMENT FOR PRIORITARIANISM

What arguments are there for prioritarianism? Let us take the value of equality for granted, so that a betterness ordering should satisfy the Pigou-Dalton condition. We are also assuming the ordering is increasing and impartial. Granting these things, the question is: what arguments support the claim that the ordering is also strongly separable? This question was raised by Parfit in the Lindley Lecture, but it was raised in economics much earlier (see Sen 1973: 39–41).

The central argument for prioritarianism in Parfit's lecture is that nonprioritarian theories are subject to the *levelling-down objection*. Compare two distributions where the second is levelled down relative to the first. This means that in the second distribution, some of the better-off people are less well off than they are in the first, but no one is better off in the second than she is in the first. Parfit claims that, though no one is better off in the second distribution, nonprioritarians must think the second is better than the first in one respect—namely, it is more equal. On the other hand, prioritarians think the second is better in no respect. Parfit claims that the former view is implausible.

Why must nonprioritarians think the levelled-down distribution is better in one respect? Take the nonprioritarian theory whose value function is the sum of the products of people's well-beings, taken in pairs. That is:

$$v(w) = w_1 w_2 + w_1 w_3 + \dots + w_1 w_n + w_2 w_3 + \dots + w_2 w_n + w_3 w_4 + \dots + w_3 w_n \quad (2)$$

This is not a well-known function. However, it is impartial and increasing, provided everyone's well-being is positive, and it satisfies the Pigou-Dalton condition. It is a well-behaved function and not obviously objectionable. It is not strongly separable, so it is nonprioritarian. The hallmark of prioritarianism is that the benefit of increasing the well-being of one person is not affected by the well-beings of other people, but in this theory it clearly is. However, a levelling down raises none of the terms in the formula (2), so on the face of it this theory does not suggest the levelled-down distribution is better in any respect.

It does not do so any more than the familiar prioritarian formula (1) does. An example of (1) is the sum of the square roots of people's well-beings:

$$v(w) = \sqrt{w_1} + \sqrt{w_2} + \dots + \sqrt{w_n} \quad (3)$$

A levelling down raises none of the terms in this formula, and Parfit tells us that for this reason it is not subject to the levelling-down objection.

Parfit argues that a nonprioritarian thinks inequality is bad, and therefore must think that any change that decreases inequality is better in one respect. But a prioritarian also thinks inequality is bad: a reduction in inequality makes the distribution better if total well-being remains constant. If a theory has a value function at all, the badness

it ascribes to inequality is easily identified. The theory can be represented by a value function of the form

$$V = (w_1 + w_2 + \dots + w_n) - I(w_1, w_2, \dots, w_n)$$

The first term in this formula is the total of people's well-being. The second is a measure of the badness of the inequality. It can be found simply by subtracting whatever value the theory ascribes to a distribution from the total of people's well-being. In the case of the square-root formula (3), the measure is

$$I(w_1, w_2, \dots, w_n) = (w_1 + w_2 + \dots + w_n) - (\sqrt{w_1} + \sqrt{w_2} + \dots + \sqrt{w_n})$$

This can be done for a prioritarian theory as easily as for a nonprioritarian one (Fleurbaey 2015: sect. 2). Levelling down decreases the measure of the badness of inequality. So here is a respect in which levelling down makes the distribution better, even for a prioritarian.

Parfit claims that a prioritarian thinks a reduction in inequality is only *instrumentally* better. The prioritarian thinks inequality is not *intrinsically* bad; it is just that shifting well-being from better-off to worse-off people, which she values intrinsically, has the effect of reducing inequality, which she does not value intrinsically. But in what sense is the badness of inequality a mere instrumental consequence of prioritarianism? It is not a causal consequence. It is not even a contingent consequence; it is a mathematically necessary feature of the prioritarian formula. The badness of inequality is intrinsic to it in the way that being the square of three is intrinsic to the number nine. Compare the pairwise-product value function (2) with the square-root function (3). Both give value to equality, because they both satisfy the Pigou-Dalton condition. How does the pairwise-product function make this value intrinsic, whereas the square-root function makes it only instrumental? I see no difference.

The upshot is that the levelling-down objection fails. It gives no support to prioritarianism. My argument for this conclusion depends on formal features of prioritarian and nonprioritarian theories, and indeed on formulae. This is the method of economics.

27.5 A GOOD ARGUMENT FOR PRIORITARIANISM

Economics also provides a quite different, powerful argument in favor of prioritarianism. It consists in a theorem first proved by John Harsanyi (1955), and subsequently proved in many different forms by other economists, including Border (1985), Broome (1990), Deschamps and Gevers (1979), Fishburn (1984), Hammond (1981), and Mongin and d'Aspremont (1998). Harsanyi discovered a clever way of getting some extra theoretical leverage on the structure of the betterness relation by investigating betterness, not just

among states of affairs, but among uncertain prospects. A prospect is a portfolio of states of affairs, each with a probability attached to it. Expected utility theory gives us information about the structure of betterness among prospects. This in turn induces some structure on the betterness relation among states of affairs. Indeed, it implies that it is strongly separable. That is Harsanyi's discovery.

Harsanyi presents his theorem in terms of preferences, but I prefer to express it in terms of betterness, as I did in Broome (1991b). The theorem rests on these three premises:

- (1) Each individual's betterness among prospects satisfies the axioms of expected utility theory.
- (2) General betterness among prospects satisfies these axioms.
- (3) One prospect is better than another if it is better than the other for one person and at least as good as the other for every person, and two prospects are equally good if they are equally good for each person.

I call the third premise "the principle of personal good." It is a translation of the Pareto principle into the terminology of betterness.

Harsanyi's theorem is that, given these three assumptions, general betterness can be represented expectationally by a value function $v()$ that is the sum of utility functions $u_1(), u_2(), \dots, u_n()$ that represent betterness for each person expectationally. That is:

$$v(x) = u_1(x) + u_2(x) + \dots + u_n(x)$$

The variable x ranges over prospects. Among prospects are those that deliver a particular outcome for sure, and these can be identified with the outcome itself. So x also ranges over outcomes. This equation therefore applies to outcomes as well as prospects.

Let us now interpret the outcomes as distributions of well-being such as $w = (w_1, w_2, \dots, w_n)$. Prospects are then portfolios of distributions, each with a probability. Harsanyi's theorem tells us that general betterness among distributions can be represented by a value function that is the sum of utility functions that represent betterness for individuals:

$$v(w) = u_1(w) + u_2(w) + \dots + u_n(w)$$

Think for a moment of just the first person. One distribution w is better for her than another w' if and only if she has more well-being in w than in w' . That is to say: if and only if w_1 is greater than w'_1 . This implies that, among distributions, betterness for the first person depends only on her own well-being w_1 . Utility for the first person represents her betterness, so it too depends only on w_1 . The same is true for each person. So we may write the equation:

$$v(w) = u_1(w_1) + u_2(w_2) + \dots + u_n(w_n) \quad (4)$$

This is an instance of the additively separable form (1) of the value function. Since additive separability implies strong separability, it follows that the betterness ordering is strongly separable. That is the powerful argument for strong separability.

We are assuming that equality is valuable, which is to say that the Pigou-Dalton condition holds. We are also assuming that the betterness ordering is increasing and impartial in people's well-being. Given these assumptions, adding strong separability gives us prioritarianism. So Harsanyi's theorem provides a powerful argument for prioritarianism.

Each function $u_i(w_i)$ shows the contribution i 's well-being makes to general value. The Pigou-Dalton condition ensures that each $u_i(w_i)$ is strictly concave in w_i . The degree of concavity determines the degree to which inequality in people's well-being is bad: the more concave, the more valuable equality is. This gives us a further consequence of the theorem. The functions $u_i(w_i)$ are defined to be expectational utility functions. This means that their degree of concavity also determines the degree to which risk to people's well-being is bad. So the badness of inequality exactly matches the badness of risk.

This consequence of Harsanyi's theorem is remarkable and surprising. It seems intuitively that the badness of inequality and the badness of risk are quite different matters. The badness of inequality is a matter of weighing together the interests of different people. This raises the issue of fairness, whereas fairness has no relevance to a single person's risk. Yet Harsanyi's theorem implies a symmetry between the badness of inequality and the badness of risk. They are both given by the concavity of the same utility functions.

Harsanyi himself drew a different conclusion. He thought his theorem constitutes an argument for utilitarianism. This is because, as I explained in Section 27.2, he assumed that the utility functions in equation (4) measure well-being cardinally: that they are affine functions of well-being. Given this assumption and also symmetry, (4) could be written:

$$v'(w) = w_1 + w_2 + \dots + w_n$$

where $v'()$ is a value function that is an affine transform of $v()$. This is a utilitarian formula for value. Utilitarianism gives no value to equality of well-being; it does not satisfy the Pigou-Dalton condition.

I explained in Section 27.2 that Harsanyi himself had no good grounds for his assumption that utility measures well-being cardinally. It appears to have arisen from a confusion over the meaning of "utility." However, some grounds can be provided, and Harsanyi's theorem itself contributes to these grounds. I shall now set them out (for more detail, see Broome 1991b: ch. 10).

Since well-being is a quantitative concept, we should have some account of the concept of a quantity of well-being. For instance, what does it mean for one quantity of well-being to be greater than another? A plausible answer is that it means that the first counts for more than the second in contexts where they are weighed against each other in determining an overall value.

One context where quantities of well-being are weighed against each other is in determining the value for a person of uncertain prospects. Here is an example. Suppose one

outcome x is better for a person than another outcome y , which is better for her than a third outcome z . And suppose y for sure is better for the person than a gamble at equal odds between x and z . The gamble is better than y in one respect and worse in another. It is better in that it might give the person x rather than y , and worse in that it might give her z rather than y . Evidently the respect in which the gamble is worse counts for more than the respect in which it is better, since the gamble is worse overall. Since it counts for more in this situation of risk, we might take the actual loss of getting z rather than y to be more than the actual benefit of getting x rather than y . That is: the difference in the person's well-being between y and z is greater than the difference in her well-being between x and y .

Since y is better than the gamble on x and z , the definition of expectational utility implies that $u(y) > \frac{1}{2}u(x) + \frac{1}{2}u(z)$. This implies that the difference $u(y) - u(z)$ in utility between y and z is greater than the difference $u(x) - u(y)$ in utility between x and y . The utility differences therefore match the putative differences in well-being that I have just described. By generalizing over many gambles, we may conclude that expectational utility measures well-being cardinally.

This is a weak argument for the conclusion that expectational utility measures well-being cardinally. It simply ignores the possibility in the example that the gamble is worse than the certainty of y , not because of the differences in quantities of well-being as I described them, but because risk to well-being is itself a bad thing.

However, the case for this analysis of quantities of well-being is much strengthened by Harsanyi's theorem. His theorem tells us that people's expectational utilities, added up, determine the relative values of distributions of well-being. This means that utility specifies, not only how differences of well-being count intrapersonally in determining the values of uncertain prospects, but also how they count interpersonally in determining the values of distributions. This makes it much more plausible that utility measures well-being cardinally. It gives strength to Harsanyi's claim that his theorem supports utilitarianism.

Still, we do not have to accept this claim of Harsanyi's. Prioritarians do not accept it. However, it leaves them with the responsibility of providing some alternative analysis of quantities of well-being (see Broome 1991b: ch. 11; McCarthy 2006). Given Harsanyi's theorem, the issue between prioritarianism, which values equality in well-being, and utilitarianism, which does not, comes down to a question about the measurement of well-being.

27.6 THE DANGERS OF IGNORING ECONOMIC THEORY

Each of the three premises of Harsanyi's theorem is *prima facie* very plausible, so the theorem provides a powerful argument for prioritarianism. Given the failure of the levelling-down argument, it is by far the strongest argument for prioritarianism.

The same theorem also sets a challenge for prioritarianism. Prioritarianism has to be held apart from utilitarianism by providing some measure of well-being distinct from its contribution to general value. Expectational utility is a natural measure to pick, but Harsanyi's theorem shows that we cannot use that measure without abandoning the value of equality. That means abandoning prioritarianism.

So Harsanyi's theorem sets the agenda for the debate about the value of equality. If you value equality but do not want to be a prioritarian, you must explain which of the theorem's three assumptions you deny and why you deny it. If you want to be a prioritarian, you must explain how you measure well-being in a way that allows you to value equality of well-being.

The literature of economics contains a considerable body of writing that responds to Harsanyi's theorem, including Coulhon and Mongin (1986), Diamond (1967), Mongin and d'Aspremont (1998), Sen (1976, 1977), and Weymark (1991). So does the literature of philosophy, including Broome (1991b), Greaves (2015), McCarthy (2006, 2008), and Rabinowicz (2002). Yet within philosophy there is also a large body of literature about prioritarianism that takes little notice of this work or its conclusions. Because it does not build on previous knowledge, this latter body of literature is badly placed to contribute to our understanding of the value of equality.

An example of it is "Why It Matters That Some Are Worse Off than Others" by Michael Otsuka and Alex Voorhoeve (2009). I pick this paper as an example because it has been highly influential. It was published in the leading journal *Philosophy and Public Affairs*, and subsequently a whole issue of *Utilitas* was devoted to comments on it. Yet this paper made no reference to Harsanyi's theorem, and nor did any of the comments in *Utilitas*.

Otsuka and Voorhoeve argue against prioritarianism. Their main argument is that different considerations are at issue when we weigh together the well-beings of different people from those that are at issue when we weigh well-being for a single person. We weigh well-being intrapersonally when we balance a chance of a gain to a person against the chance of a loss to the same person. We weigh well-beings interpersonally when we balance a gain to someone against a loss to someone else. Interpersonal weighings raise the issue of fairness, whereas intrapersonal weighings do not. The paper supports this point by means of an example in which it seems clear that interpersonal and intrapersonal weighings should be done differently. The authors say that failing to do so is failing to recognize the separateness of persons.

They offer this as an argument against prioritarianism. But it is most directly an argument against the conclusion of Harsanyi's theorem, although the authors do not identify it as such. Harsanyi's theorem asserts a symmetry between interpersonal weighing and interpersonal weighing. They deny this symmetry.

Their objection to symmetry is an old one. It was raised by Peter Diamond (1967) in response to Harsanyi and has been a subject of discussion ever since. Otsuka and Voorhoeve support it with an example but, given the long history, an example does not advance the argument much. On the one hand, the example elicits an intuition that the theorem's conclusion is wrong. On the other hand, there is the attraction of the theorem's three plausible premises. We have to navigate between these conflicting

considerations. If we are to accept the intuitive objection to symmetry, we have to identify what is wrong with at least one of the three premises.

The literature has canvassed the options. Different authors have taken different stances. Diamond himself objected to the second premise that general betterness conforms to expected utility theory. He provided a direct counterexample to this premise. Accepting Diamond's objection, David McCarthy (2006) also gives up the second premise and develops a version of prioritarianism—*ex ante prioritarianism*—that he claims to be defensible. On the other hand, Wlodek Rabinowicz (2002) recommends giving up the third premise—the principle of personal good—and he provides a different version of prioritarianism on that basis. (He distinguishes two sorts of intrapersonal weighing of well-beings in determining betterness among prospects: weighing for determining betterness for the person and weighing for determining general betterness. He denies that the first sort is symmetrical with interpersonal weighing, but accepts that the second sort is.) McCarthy and Rabinowicz have each worked out their own coherent prioritarian theory.

Otsuka and Voorhoeve do not try to navigate the conflict between their intuition and the premises of Harsanyi's theorem. They do not identify which premise they reject. Indeed, they apparently affirm all three at different points in their paper. They affirm the first (expected utility theory for individual betterness) and second (expected utility theory for general betterness) explicitly (Otsuka and Voorhoeve 2009: 172n3 and 195 respectively). They do not explicitly affirm the third (the principle of personal good). However, they affirm it implicitly in responding to Rabinowicz's version of prioritarianism. Take two prospects that are equally good for everyone apart from one person. If I understand the authors correctly, they assume that one of these prospects is better than the other if and only if it is better for that one person (2009: 178). By easy steps, this assumption implies the principle of personal good.

I say that the authors make these assumptions only "apparently" because they do not use the language of betterness. In the situation where two prospects are equally good for everyone apart from one person (and there is a question of what medical treatment to provide for that person), they say "it is reasonable to provide [the person] with a treatment that maximizes the expected increase in her utility." It appears later on the same page that they use "reasonable" in such a way that doing anything else counts as unreasonable. I therefore take them to mean that the person ought to be provided with a treatment that maximizes the expected increase in her utility. That seems to imply that this treatment gives the best prospect.

So there are *prima facie* grounds for thinking that Otsuka and Voorhoeve have fallen into a trap of inconsistency. They apparently affirm the premises of Harsanyi's theorem but deny its conclusion. This is a trap they set for themselves by choosing to ignore Harsanyi's work. If the inconsistency is real, it vitiates their argument. At least the authors should have explained how they avoid the threat of inconsistency.

The paper appears to contain a further inconsistency. In responding to McCarthy's version of prioritarianism, the authors in effect deny the principle of personal good (Otsuka and Voorhoeve 2009: 197–8). Apparently, they affirm this principle and later

deny it. They make various separate affirmations through the paper, but seem not to have checked them for consistency. To be sure they are consistent, they would need to set out a theory of their own that conforms to them, as Rabinowicz and McCarthy do but they do not.

27.7 THE USES OF MATHEMATICS

How can the authors, referees, editors, and many readers of this influential paper not have checked its consistency? I am sorry to say that the practice of moral philosophers is not always as analytically tight as it should be. Often we shun formal language, we avoid symbols, and we do not use mathematics. This is appropriate for wide swathes of moral philosophy, including those that are concerned with personal relationships among small numbers of people. But it is not appropriate in areas of moral philosophy that involve large numbers of people and have a quantitative dimension. The value of equality is one of those.

In areas of this sort we need mathematical formulae to keep our thinking accurate. Moreover, we can harness the power of mathematical methods. Harsanyi's theorem illustrates their power. Its premises contain no mention of addition, yet its conclusion is that the value function has an additively separable structure. No amount of verbal discussion or probing of examples could arrive at this strong and surprising result. Yet the result can be used to provide powerful support for the philosophical theories of utilitarianism and prioritarianism.

This is another good lesson we can draw from economics: sometimes we need mathematics. Economics became mathematical in the 1940s and 1950s. After that time, it was impossible to become an academic economist without at least being able to understand mathematical formulae. Moral philosophy does not need such a radical revolution in its methods, but we must recognize that some topics within moral philosophy cannot be studied properly without mathematics. Furthermore, many moral philosophers cannot benefit from the discoveries of economists that I have been commending because, without mathematical understanding, they cannot read the papers.

A lack of mathematical understanding also leads to errors at a more mundane level, where no fancy theorems are required. Simple mathematical sensitivity would often be beneficial. Here is an example.

Let q be a variable ranging over qualities and let t be a variable ranging over quantities. The vector (q, t) , which I shall call an "item," denotes a particular quantity of a particular quality. For instance, it may denote a life of quality q that lasts for a time t . Or it may denote a population of t people, each with a lifetime well-being of q . Or it may denote an episode of pain of quality q lasting for a period of time t . And so on.

Assume the items within some domain, such as the domain of episodes of pain or the domain of lives, are completely ordered by their goodness. This means in particular that all the items having some standard quantity T are ordered. We can treat the ordering of

those particular items as an ordering of the qualities q themselves. That is, we can define q to be better than q' if and only if (q, T) is better than (q', T) . Let q_h and q_l be two qualities such that q_h is better than q_l . For example, q_h might be a high quality of life and q_l a mediocre one. Or q_l might be an excruciating degree of pain and q_h the pain of a slight headache.

Suppose it is true that:

Premise. For any item (q, t) , there is a quality q' worse than q and a quantity t' such that (q', t') is better than (q, t) .

That is, for some diminution in quality, a sufficient increase in quantity more than cancels out the diminution and leads to a result that is better overall.

Even though betterness is a transitive relation, it does not follow from the Premise that:

Conclusion. There is a quantity t_l such that (q_l, t_l) is better than (q_h, T) .

That is, it does not follow that a sufficient increase in quantity more than cancels out a diminution in quality all the way from q_h to q_l . Even given the Premise, it may be that an item with quality q_h and quantity T is better than any item with the lower quality q_l , whatever its quantity.

The premise implies that, starting from (q_h, T) there is a sequence of qualities $q_i, q_j, \dots$ that are progressively worse and worse—so they are all worse than q_h —but such that with sufficient increases in quantity, the items $(q_i, t_i), (q_j, t_j)$ are progressively better and better, so they are all better than (q_h, T) . But the sequence $q_i, q_j, \dots$ may never get as low in the ordering of qualities as q_l .

The same is true if we reverse the direction of change. This is appropriate if the domain consists of bad things such as pains, so that increasing quantity makes an item worse.

Reverse Premise. For any vector (q, t) , there is a quality q' better than q and a quantity t' such that (q', t') is worse than (q, t) .

It does not follow that:

Reverse Conclusion. There is a quantity t_h such that (q_h, t_h) is worse than (q_l, T) .

It will be obvious to anyone with a mathematical sensitivity that the Conclusion does not follow from the Premise, or the Reverse Conclusion from the Reverse Premise. But several philosophers have assumed the opposite, and drawn extravagant conclusions. In one of his papers, Larry Temkin (1966) worked with the domain of pains. He accepted the Reverse Premise and assumed it implied the Reverse Conclusion, given that betterness is transitive. He drew the conclusion that betterness is not transitive. Temkin was reproved for his mistake by Ken Binmore and Alex Voorhoeve (2003), and I believe he

has avoided it ever since. Nevertheless, Dale Dorsey (2009) later made exactly the same mistake. He concluded that the betterness ordering is discontinuous.

Derek Parfit (1984) dealt with a domain of populations. He chose a very high value for q_h and a value of q_l that he took to make the Conclusion “repugnant.” He presented persuasive arguments for the Premise. Then he assumed that the Conclusion follows from the Premise. This presented him with a problem (Parfit 1984: 430, 435–6). But since the Conclusion does not follow from the Premise, this problem did not really arise. Whatever problem there is with the Repugnant Conclusion, it is not a problem for the Premise. (The Repugnant Conclusion is indeed implied by a version of utilitarianism.)

In a much later paper, in dealing with the domain of populations, Parfit (2016) did not assume the Conclusion follows from the Premise alone. He added a further premise that I shall mention below. But in the domain of lives of different lengths, he continued implicitly to assume the Conclusion follows from the Premise alone (Parfit 2016: 119).

A counterexample helps to show the error in this assumption more clearly. Suppose q and t each have a numerical measure. Suppose the betterness ordering can be represented by a function $v(q, t)$, and let this function take the specific form

$$v(q, t) = qt/(t + 1). \quad (5)$$

Figure 27.1 shows this function’s contours of constant goodness. These could be called indifference curves for goodness. (Another lesson from economics is that indifference curves are a good aid for understanding the structure of value.) With this function, the Premise is true and the Conclusion false.

The specific form of the function does not matter; other formulae would do instead. What does matter is that the contours slope downwards everywhere and that they approach horizontal asymptotes. Their downward slope explains why the Premise is true: a sufficiently small diminution in quality (a downward movement) can always be

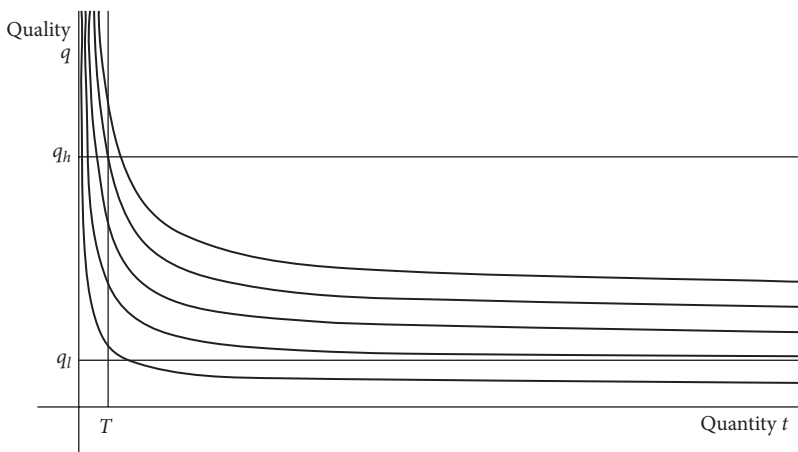


FIGURE 27.1 Contours of equal goodness.

compensated for by a sufficiently large increase in quantity (a rightward movement). The asymptotes explain why the Conclusion is false. No item with a quality below the asymptote of a particular contour, however great the quantity, can be as good as any item on or above the contour.

The ordering has the property that James Griffin (1986: 85) calls “discontinuity” and defines as “So long as we have enough of *B*, any amount of *A* outranks any further amount of *B*, or ... enough of *A* outranks any amount of *B*.” (In the second disjunct read *A* as q_h and *B* as q_l .) I assume Griffin was aiming at the mathematical notion of discontinuity, but he did not hit it. The ordering in this example satisfies Griffin’s definition but it is continuous in the mathematical sense. Discontinuous orderings can be intractable, but this ordering is not. Indeed, it is very tractable. I commend an ordering with horizontal asymptotes to philosophers who are struggling with population ethics. It offers a way around some of their problems.

Interestingly, the formal structure of Gustaf Arrhenius’s forthcoming book *Population Ethics* ensures that the Conclusion actually does follow from the Premise. His “First impossibility theorem” states as much. In his structure, quality comes in discrete amounts. The consequence is that the contours of the betterness ordering cannot have horizontal asymptotes while always sloping downwards. This is merely an artefact of his idiosyncratic assumption that quality comes in discrete amounts.

Derek Parfit (2016: 162) makes an objection to a betterness relation whose contours have horizontal asymptotes. He says it implies that the existence of more and more people has “diminishing marginal value.” Adding each new person to the world has less and less value, the more people there already are. Parfit thinks this cannot be correct, and he consequently rejects this counterexample to the claim that the Premise implies the Conclusion.

What matters is the form of the betterness relation, not the value function that represents this relation. A value function is arbitrary to a large extent. A value function for the betterness ordering in Figure 27.1 can be created by giving a value to each contour. Any values at all can be assigned to the contours, so long as a higher contour always gets a higher value than a lower one. Any valuation of contours that sticks to this rule gives a representation. This means the scale of value is arbitrary, so that the marginal value of quantity is also arbitrary. The value function must have the right contours, but it is arbitrary apart from that.

Despite what Parfit supposes, a betterness ordering whose contours have horizontal asymptotes does not imply that quantity has diminishing marginal value. The value function (5) has that implication, but it is only one of the many different value functions that represent the same ordering. Others do not imply diminishing marginal value.

It is true that, when an ordering has contours with horizontal asymptotes, any representation of it has the feature that value is not everywhere linear in quantity. Quantity cannot always have constant marginal value. Indeed, in any representation, as quantity gets bigger and bigger, it eventually has to have diminishing marginal value. This much of Parfit’s supposition is true. But it has little significance, since the point at which the

marginal value of quantity begins to decline can be extended as far away as anyone might choose.

For example, think of items (q, t) denoting populations of size t at the low level of well-being q . There is a representation that makes the value of (q, t) increase in proportion to t until t reaches, say, a trillion trillion. That is to say, the marginal value of adding a person at level q is constant until the population reaches a trillion trillion. If that is not enough, we can go further.

To see how this is possible, start by assigning the value of t to the contour that passes through (q, t) , and do this for every t up to some large number. To do this in Figure 27.1, take the horizontal line at level q , and assign to every contour that cuts this line the value of t at the point where it cuts it. Algebraically, this is done by transforming the value function $v()$ in (5) to $v'()$ where:

$$v'(q, t) = v(q, t) / \{q_t - v(q, t)\}$$

This means that

$$v'(q, t) = qt / \{q_t(t + 1) - qt\}$$

If you were to continue this method of assigning value for every t out to infinity, only contours that cut the line would get a value. There are many contours that do not cut that line. To assign them a value, at some point you will have to change your valuing method. But you can continue with the original method as far as you like. (If you were willing to accept values that are transfinite numbers, you could even take t out to infinity.)

So Parfit's objection to the counterexample is mistaken. To reveal the mistake, I applied a lesson from economics that I mentioned in Section 27.3: I adopted the economists' practice of concentrating on the form of the betterness relation rather than on a value function that represents it.

27.8 CONCLUSION

The main lesson of Section 27.7 is perhaps the most important of the lessons economics can teach moral philosophy, because it is a precondition for learning the others. Moral philosophers should learn that some of the subjects that interest them demand a little mathematics. If you are not interested in mathematics, you should avoid these subjects.

One benefit of knowing some mathematics is that it will make the previous work of economists available to you. Economists have done excellent work on some of the topics that also interest philosophers. There are many particular lessons to be learnt from their results. Philosophy can build on them, but it must absorb them first. I mentioned two examples.

I described how far the moral philosophy of equality can go astray if it ignores the long history of the analysis of equality in economics. In particular, long before philosophers came to the subject, economists had already investigated the foundations of prioritarianism. They made discoveries that depend on remarkable theorems about additivity. These discoveries could not have been made except by mathematical methods. The philosophy of prioritarianism cannot be successful if it ignores them.

There is also a long history of population ethics in economics, driven by the work of Charles Blackorby, Walter Bossert, and David Donaldson. These authors have developed a complete axiology for population (Blackorby et al. 2005). But for the example of population ethics, I concentrated on just one mathematical error that has been damaging. This error also has ramifications outside population ethics, and it has led to some serious mistakes.

In presenting my examples I mentioned some other lessons in passing. An important one is that, in the theory of value, it is best to think first of the structure of the betterness ordering rather than of value itself. Value merely represents betterness. As I put it in Broome (1999: 9–11), the lesson is: “Think comparatively.” Attend to what is better than what rather than to what is good. Comparative thinking is deeply embedded in economics but not in philosophy.

Finally, I strongly recommend avoiding the economist’s practice of using “utility” to denote a person’s good.

ACKNOWLEDGMENTS

I am very grateful to Wlodek Rabinowicz and David McCarthy for detailed advice about this paper. Research for the paper was supported by ARC Discovery Grants DP140102468 and DP180100355.

REFERENCES

- Arrhenius, Gustaf. Forthcoming. *Population Ethics*.
- Arrow, Kenneth J. 1951. *Social Choice and Individual Values*. New Haven, CT: Yale University Press.
- Atkinson, Anthony B., and Joseph E. Stiglitz. 1980. *Lectures on Public Economics*. New York: McGraw-Hill.
- Bernoulli, Daniel. 1738. “Specimen theoriae novae de mensura sortis,” translated by Louise Sommer and published in 1954 as “Exposition of a New Theory on the Measurement of Risk.” *Econometrica* 22: 23–36.
- Blackorby, Charles, Walter Bossert, and David Donaldson. 2005. *Population Issues in Social Choice Theory, Welfare Economics and Ethics*. Cambridge: Cambridge University Press.
- Border, Kim C. 1985. “More on Harsanyi’s Cardinal Welfare Theorem.” *Social Choice and Welfare* 2: 279–81.
- Broome, John. 1990. “Bolker-Jeffrey Expected Utility Theory and Axiomatic Utilitarianism.” *Review of Economic Studies* 57: 477–502.

- Broome, John. 1991a. "Utility." *Economics and Philosophy* 7: 1–12.
- Broome, John. 1991b. *Weighing Goods: Equality, Uncertainty and Time*. Oxford: Blackwell.
- Broome, John. 1999. *Ethics Out of Economics*. Cambridge: Cambridge University Press.
- Binmore, Ken, and Alex Voorhoeve. 2003. "Defending Transitivity Against Zeno's Paradox." *Philosophy and Public Affairs* 31: 272–9.
- Coulhon, T, and P. Mongin. 1986. "Social Choice Theory in the Case of von Neumann-Morgenstern Utilities." *Social Choice and Welfare* 6: 175–87.
- Deaton, Angus, and John Muellbauer. 1980. *Economics and Consumer Behavior*. Cambridge: Cambridge University Press.
- Debreu, Gerard. 1954. "Representation of a Preference Ordering by a Numerical Function." In R.M. Thrall, C.H. Coombs, and R.L. Davis (eds.), *Decision Processes* (New York: John Wiley & Sons), pp. 159–65.
- Debreu, Gerard. 1959. *Theory of Value: An Axiomatic Analysis of Economic Equilibrium*. New York: John Wiley & Sons.
- Deschamps, Robert, and Louis Gevers. 1979. "Separability, Risk-Bearing and Social Welfare Judgments." In Jean-Jacques Laffont (ed.), *Aggregation and Revelation of Preferences* (Amsterdam: Elsevier), pp. 145–60.
- Diamond, Peter A. 1967. "Cardinal Welfare, Individualistic Ethics, and Interpersonal Comparison of Utility: Comment." *Journal of Political Economy* 75: 765–6.
- Dorsey, Dale. 2009. "Headaches, Lives and Value." *Utilitas* 21: 36–58.
- Fishburn, Peter C. 1984. "On Harsanyi's Utilitarian Cardinal Welfare Theorem." *Theory and Decision* 17: 21–8.
- Fleurbaey, Marc. 2015. "Equality Versus Priority: How Relevant Is the Distinction?" *Economics and Philosophy* 31: 203–17.
- Greaves, Hilary. 2015. "Antiprioritarianism." *Utilitas* 27: 1–42.
- Griffin, James. 1986. *Well-Being: Its Meaning, Measurement and Moral Importance*. Oxford: Oxford University Press.
- Hammond, Peter J. 1981. "Ex-Ante and Ex-Post Welfare Optimality under Uncertainty." *Economica* 48: 235–50.
- Harsanyi, John C. 1955. "Cardinal Welfare, Individualistic Ethics, and Interpersonal Comparisons of Utility." *Journal of Political Economy* 63: 309–21.
- Harsanyi, John C. 1975. "Can the Maximin Principle Serve as a Basis for Morality? A Critique of John Rawls's Theory." *American Political Science Review* 69: 594–606.
- Harsanyi, John C. 1977. *Rational Behavior and Bargaining Equilibrium in Games and Social Situations*. Cambridge: Cambridge University Press.
- McCarthy, David. 2006. "Utilitarianism and Prioritarianism I." *Economics and Philosophy* 22: 1–29.
- McCarthy, David. 2008. "Utilitarianism and Prioritarianism II." *Economics and Philosophy* 24: 1–33.
- Mongin, Philippe, and Claude d'Aspremont. 1998. "Utility Theory and Ethics." In Salvador Barbera, Peter Hammond, and Christian Seidl (eds.), *Handbook of Utility Theory, Volume 1: Principles* (Dordrecht: Kluwer), pp. 371–480.
- Otsuka, Michael, and Alex Voorhoeve. 2009. "Why It Matters That Some Are Worse Off than Others: An Argument Against the Priority View." *Philosophy and Public Affairs* 37: 172–99.
- Otsuka, Michael, and Alex Voorhoeve. 2018. "Equality Versus Priority." In Serena Olsaretti (ed.), *The Oxford Handbook of Distributive Justice* (Oxford: Oxford University Press), pp. 65–85.

- Parfit, Derek. 1984. *Reasons and Persons*. Oxford: Oxford University Press.
- Parfit, Derek. 1991. "Equality or Priority?" The Lindley Lecture, University of Kansas. Available at <https://kuscholarworks.ku.edu/handle/1808/12405>.
- Parfit, Derek. 2016. "Can We Avoid the Repugnant Conclusion?" *Theoria* 82: 110–27.
- Rabinowicz, Wlodek. 2002. "Prioritarianism for Prospects." *Utilitas* 14: 2–21.
- Ramsey, Frank. 1931. "Truth and Probability." In *Foundations of Mathematics and Other Logical Papers*, R. B. Braithwaite (ed.) (Abingdon, UK: Routledge and Kegan-Paul), pp. 156–98.
- Rawls, John. 1971. *A Theory of Justice*. Cambridge, MA: Harvard University Press.
- Sen, Amartya. 1970. *Collective Choice and Social Welfare*. San Francisco, CA: Holden-Day.
- Sen, Amartya. 1973. *On Economic Inequality*. Oxford: Oxford University Press.
- Sen, Amartya. 1976. "Welfare Inequalities and Rawlsian Axiomatics." *Theory and Decision* 7: 243–62.
- Sen, Amartya. 1977. "Non-Linear Social Welfare Functions: A Reply to Professor Harsanyi." In Robert E. Butts and Jaakko Hintikka (eds.), *Foundational Problems in the Special Sciences* (Dordrecht: D. Reidel), pp. 297–302.
- Temkin, Larry S. 1966. "A Continuum Argument for Intransitivity" *Philosophy and Public Affairs* 25: 175–210.
- Tintner, Gerhard. 1942. "A Contribution to the Non-Static Theory of Choice." *Quarterly Journal of Economics* 56: 274–306.
- von Neumann, John, and Oskar Morgenstern. 1944. *Theory of Games and Economic Behavior*. Princeton, NJ: Princeton University Press.
- Weymark, John A. 1991. "A Reconsideration of the Harsanyi-Sen Debate on Utilitarianism." In Jon Elster and John Roemer (eds.), *Interpersonal Comparisons of Well-Being* (Cambridge: Cambridge University Press), pp. 255–320.

INDEX

Note: Page numbers in **bold** refer to authors' main entries; page numbers in *bold italics* refer to figures.

A

- Abeler, Johannes 185, 185n2, 187, 188, 203
 absolutist and moderate deontology 526–7
 Acemoglu, Daron 575, 578
 achieved functionings 98, 99
 act-consequentialism 532
 act types 142, 144, 145
 adaptive complex systems, agent-based 151–2
 Addabbo, Tindara 258, 265
 Adler, Matthew D. xv, 6, **399–420**, 426, 434, 441, 465, 471, 478n3, 479, 479n4, 480, 488, 493, 524n1, 532, 535, 538
 Adler, Moshe 14, 27, 30
 Adorno, Theodor W. 283–4, 288
 adversarial behavior 351
 advocacy 140, 237, 280, 448
 affiliation 104, 161, 393,
 Affordable Care Act 507
 Africa 85
 African Americans 196, 563, 570, 571, 577, 578
 age 128, 142, 170, 252, 567, 569
 agency 35, 36, 37, 40, 43, 45, 48, 97, 221, 252, 262, 264, 302, 307, 312, 477, 483, 514
 achievement 222, 224
 freedom 222, 223, 224
 agent-based adaptive complex systems 151–2
 aggregation theorem (Harsanyi) 426, 451
 AIDS 460
 Aiken, Alexander M. 448, 471
 Ainslie, George 47, 50
 Akerlof, George A. 14, 30, 62n17, 73
 Al-Najjar, Nabil I. 436, 441
 al-Nowaihi, Ali xv, 5, **181–203**, 204
 Aldrich, Katie 493
 Alesina, Alberto 270, 274, 278, 282, 285, 286, 287
 Alexander, J. McKenzie 151, 153, 155
 Alexander, Larry 525, 527, 530, 538, 539
 Alexander, Richard 126, 132
 alienation 23, 24, 25, 29, 305, 306, 320, 321, 322
 Alm, James 186, 203
 Altonji, Joseph 573, 578
 altruism 2, 21, 23, 36, 56, 59, 63, 64, 90, 118, 119, 120, 125, 126, 132, 151n14, 184, 197, 197n5, 332
 Alvey, James 36, 50
 ambiguity aversion 426, 435–7, 440, 441
 America(ns) 28, 149, 174, 233, 318, 341, 386, 393, 504, 534, 566, 577
 American Economic Association (AEA) 87, 88, 93, 474
 Americans with Disabilities Act (ADA) 575
 Amiel, Yoram 439, 442
 Anderson, Elizabeth 36, 40, 50, 112, 252, 265, 311n16, 313, 393, 395
 Andreoni, James 63, 73, 128n15, 132
 androcentric ethics 248, 254–5
 beyond 248–51
 androcentrism 248, 249,
 Angel, James J. 367, 376
 anger 128, 198
 Angner, Erik 37, 50, 104, 112, 478, 493
 Angrist, Joshua 575, 578
 Annas, Julia 37, 38, 39, 48, 48n7, 50
 anonymity 417, 418, 419, 451, 452
 and morality 197
 anthropology 58, 118, 139, 256n8, 270, 272, 274, 275, 276, 278, 281, 282, 360
 anti-discrimination laws 524, 532, 534, 535, 536, 537, 562, 575, 576
 antitrust law 71, 391, 395
 Apel, Karl-Otto 249

- Applbaum, Arthur 349, 350, 351, 351n3, 356
 applied economics 238, 399, 403, 478n3
 ethics in 497–578
 applied ethics 43, 97, 557
 approval 15, 18, 28, 120, 127, 164, 166, 168, 169,
 170, 171, 301
 Aquinas, Thomas 365, 376
 Argyle, Michael 29n20, 30
 Ariely, Dan 200, 203, 206
 Aristotle 4, 34, 36n2, 37, 37n3, 38, 39, 44, 46,
 47, 79, 259, 272, 280, 363, 364, 365,
 376, 550
 Arius Didymus 46, 47
 Armantier, Olivier 186, 203
 Armed Forces Qualifying Test (AFQT) 573
 Armendáriz, Beatriz 372, 373, 376
 Arn, Christof 258, 265
 Arneson, Richard 97, 112
 Arnold, Denis G. 373, 376
 Arnold, Matthew 275, 287, 323n6, 335
 Arrhenius, Gustaf 602, 604
 Arrow, Kenneth J. 86, 92, 93, 130, 132, 423,
 442, 504, 518, 587, 588, 604
 Arrow-Debreu extension of general
 equilibrium theory 423
 Arthur, W. Brian 147, 155
 arts 270, 276, 277, 298
 role in society 86–7
ashmot (guilt) offerings 189
 Asheim, Geir B. 441, 442
 Ashraf, Nava 14n4, 30
 Ashta, Arvind 373, 377
 asset 361–2
 asset bubble 362
 assurance 169
 asymmetric information 88
 Atkinson, Anthony B. 452n3, 471, 589, 604
 atonement 189
 Augustine 37
 Aurelius, Marcus 49, 50
 austerity 332
 Australia 28
 Austria 186
 Austrian School of Economics
 economics and ethics within 229–45
 social cooperation 240–4
 authoritarian regimes 333
 authority 130, 130n17, 298, 392, 482
 obedience to 123, 124
 autonomy 55, 56, 56n5, 57, 60, 67, 71, 72, 79,
 83, 104, 111, 220, 233, 250, 263, 368,
 393, 395, 404, 499, 511, 524, 536
 Autor, David 567, 578
 Auyang, Sunny Y. 147, 155
 Avineri, Shlomo 322, 335
 avoidable vs. unavoidable harms 476,
 486, 487
 avoiding conflicts of interest 368–9
 Ayub, Muhammad 365, 376
 Axelrod, Robert 151, 156
 Azar, Ofer H. 186, 203
- B**
 Backhouse, Roger E. 37, 50, 54n1, 73
 “bad man” (Holmes) 65
 Badgett, M.V. Lee 259, 260, 265
 Badhwar, Neera K. 29n20, 30
 Baier, Andrea 261, 265
 Bakan, Joel 339, 341, 343n1, 356
 Baker, Jennifer A. xv, 4, 34–50, 532n4, 539
 Balafoutas, Loukas 194, 203
 Balkin, Jack M. 530, 539
 Bangladesh 189, 372
 banks 195, 323, 340, 361, 365, 368, 371, 372
 Baradaran, Mehrsa 365, 376
 Barber, Benjamin 309n9, 313
 bargaining 200, 201, 236, 306, 389, 390, 395,
 479n4, 490, 512, 534, 560, 563, 573
 Barkan, Rachel 188, 203
 Barone, Enrico 325
 barter 22, 45, 243, 303, 319n3, 352, 359
 economy 359, 360
 Bartling, Bjorn 193, 203
 Bastiat, Frédéric 240, 333n14, 335
 Basu, Kaushik 110, 112
 Bateman, Milford 373, 376
 Battigalli, Pierpaolo 199, 203
 Bauhardt, Christine 253, 265
 Baumeister, Roy F. 62, 73
 Baumol, William J. 68, 73, 276, 287, 331, 335
 Baumol disease 277
 Baurmann, Michael 37, 39, 50
 Beauvoir, Simone de 250, 265
 Bechuk, Lucian 343, 356

- Becker, Gary S. 2, 7, 35n1, 50, 64, 86, 93, 131, 132, 568, 578
- Becker, Lawrence 39, 46n6, 50, 73
- Beecher, Catherine E. 259, 265
- behavioral economics 3, 5, 14, 37, 176, 181–207
- behavioral opportunism 162, 163
- beliefs and morality: psychological
game theory 198–200
- Ben-Ner, Avner 128n15, 132
- benefit corporations 64n19, 342
- Benería, Lourdes 256, 265
- benevolence 14, 17, 18, 21, 26, 28, 37, 39, 79, 239, 241, 300, 302, 312, 510
limits of 21–3
- Bennett, Jonathan 525, 539
- Benhabib, Seyla 249, 265
- Bennholdt-Thomsen, Veronika 260, 265
- Bentham, Jeremy 3, 79, 82, 232, 245, 365, 376, 418
- Bergh, Jeroen C.J.M. van den 399, 421
- Bergstrom, Theodore C. 2, 7
- Berlin, Isaiah 92, 93, 107, 112, 318n4, 335, 511n8, 518
- Bernoulli, Daniel 586–7, 604
- Bernoulli's Hypothesis 586–7, 588
- Berry, Christopher J. 14, 19n9, 30
- Bertrand, Marianne 196, 206, 570, 571, 578
- Better Business Bureau 81
- between-group selection 150
- Betzler, Monika 59n11, 73
- Beugelsdijk, Sjoerd 270, 274n2, 276n3, 278n4, 285, 287
- Beveridge, Sir William 237, 245
- Bicchieri, Cristina 140–1, 142, 155, 156
- bilateral setting 201
- Bill of Rights (US) 69
- Binder, Christina 100, 112
- Binder, Constanze xv–xvi, 4, 96–111, 112, 256n9
- Binmore, Kenneth 119n5, 132, 139, 156, 161, 177, 600, 605
- binomial tests 184
- biology 125, 127, 128, 130, 131, 139, 250, 515
- bioethics 97
- Birnie, Arthur 365, 376
- Black, Dan A. 574, 578
- black markets 352
- Blackorby, Charles 404, 421, 424, 428, 442, 604
- blame 19n10, 20n11, 190, 193, 199, 312, 367, 505
- Blanc, Louis 318
- Blanchet, Didier xvii
- Blank, Rebecca 573, 578
- Blaug, Mark 214, 227
- Blinder-Oaxaca decomposition 572
- Block, Walter 234, 245
- blood donations 201
- Blount, Sally 193, 203
- Boadway, Robin 400, 402, 404, 412, 421, 479, 479n4, 493
- Boardman, Anthony E. 402, 421
- Boatright, John R. 368, 369, 370, 376
- Bodo rules 142–3
- Boehm, Christopher 142n3, 143n7, 151n13, 156
- Boettke, Peter J. xvi, 2, 5, 7, 229–45
- Bogle, John C. 366, 376
- Bok, Sissela 122n7, 132
- Boltanski, Luc 272, 287, 338, 356
- Boly, Amandou 186, 203
- Bommier, Antoine 430, 431, 442
- Bonikowski, Bart 570
- bookshelf metaphor 165–70
- Boorse, Christopher 500n2, 518
- Border, Kim C. 593, 604
- Borna, Shaheen 366, 376
- Bosanquet, Bernard 144n9, 156
- Bossert, Walter 404, 421, 604
- bottom-up social self-organization 155
- Botzen, W.J. Wouter 399, 421
- Boumans, Marcel xvi
- Bourdieu, Pierre 284, 287
- bourgeois virtues 29, 38, 270, 280, 283
- Bouris, Erica 481, 493
- Bovens, Luc xvi, 6, 446–71, 472
- Bowie, Norman 364, 376
- Bowles, Samuel 2, 7, 51, 81, 89, 93, 94, 118nn2,3, 120, 125, 128n15, 131, 132, 142n5, 148, 150, 151nn13,14, 156, 200, 203, 381, 395
- Boyd, Robert 118n3, 124, 127, 128, 128n15, 132, 134, 142n4, 143n7, 148, 150, 151, 151n13, 158, 163n5, 177
- Bradley, Richard 437, 442
- Brandt, Richard B. 143, 156

Brecht, Bertolt 83, 93
 Brennan, Geoffrey 37, 39, 50, 535n5, 539
 Brennan, Jason 307n9, 308n11, 313, 548, 557, 558
 Brennan, Samantha 522, 526, 527, 530, 539
 Brexit 275
 Britain 121, 230, 275, 299
 Brock, Dan W. 461n7, 472, 507, 518
 Broome, John xvi, 7, 66n24, 424, 425, 442, 501, 518, 583–604, 605
 Brown, Abbey 493
 Brown, Claire 279
 Brown, Donald E. 122n7, 132
 Brown, Marvin T. 262, 265
 Bruce, Neil 404, 412, 421, 479, 493
 Bruin, Boudewijn de 359n1, 368, 369, 376
 Bruni, Luigino 46, 50
 Brunnermeier, Marcus K. 371, 372, 376
 Buchak, Lara 459n5, 472
 Buchanan, Allen 129, 133
 Buchanan, James McGill 66n25, 73, 238, 239, 240, 245, 480, 493
 Buddhism 79, 279
 Bullard, Robert 545n1, 558
 bullion 360
 Burke, Edmund 149, 149n11, 156
 Burns, Robert 548
 butcher-brewer-baker passage (Smith, WN) 300, 301, 302
 Butler, Judith 250, 251, 265
 Buyx, Alena 509, 520
 Bykvist, Krister 401, 421
 Byskov, Morten F. 102, 112

C

Caesarian section 6, 449–50
 and vaginal birth 458, 459
 Calabresi, Guido 521, 523, 533, 539
 calculative trust 168
 Caldwell, Bruce 334
 Callon, Michael 276, 287
 Calvin, John 279
 Cambridge school of economics 231, 233
 Camerer, Colin F. 14n4, 30, 51, 120, 126n13, 131, 133, 182, 183, 186, 204
 Cameron, Kim 270, 281, 287
 Campbell, John 87, 93

cancer 6, 448, 449, 457, 458, 470, 471, 504
 Cannan, Edwin 231
 capabilitarian theories of justice 96, 98, 99
 capabilities 4, 86, 96, 98, 98n1, 103n2, 106n4, 128, 162, 219, 256, 256n9, 321, 401, 484, 489
 and moral values and human dignity 221–4
 capability accounts 98
 capability approach 96–111
 and assessment of economic institutions and systems
 how does it make a difference 106–9
 how should it be further developed 109–11
 in economic ethics 106–11
 and freedom 100–3
 what is it? 97–100
 capability set 98, 98n1, 99, 101, 102, 103, 104, 107, 108, 110
 capability theories 96, 98, 105, 106
 capitalism 5, 81, 83, 106, 107, 176n11, 229, 230n1, 231, 234, 240n3, 244, 253, 256n8, 260, 279, 280, 284, 306, 337, 339, 341, 355, 356, 363, 364, 366, 381, 387, 491
 and democracy 316–34
 Hayek 324–8
 Marx 320–2
 Polanyi 328–30
 postwar literature 330–4
 Schumpeter 322–4
 Tocqueville 318–20
 varieties of 330–1
 care
 ethics 64n20, 250, 251, 263
 global theory 250
 justice 250, 264, 265
 for whom? 250–1
 migration, moral harm of 251
 political theory of 250
 rights 249
 work 252, 258, 259, 260, 261, 263, 264
 Careful Economics 263
 caring 27, 82, 181, 243, 255, 562
 citizenship 251
 democracy 251
 masculinities 251

- provisioning 262–4
 feminist economics of 264
 rationality 250
 Carpenter, Jeffrey P. 128n15, 133, 195n4, 204
 Carson, Richard T. 534, 539, 541
 Carter, Ian 101, 112
 Cat, Jordi 209, 226, 227
 catastrophe avoidance 440, 464
 vs. equity 432–3
 categorical imperative (Kant) 56, 57, 57n6, 58,
 79, 249, 305, 349, 508
 Cathcart, Thomas 56n3, 73
 Catholicism 189, 279, 280, 331, 562
 CBA *see* cost-benefit analysis
 CEDAW (Convention of the Elimination of
 all Forms of Discrimination against
 Women) 249
 certainty equivalent 431, 454, 457, 459,
 460, 464
 Chambers, Christopher P. 428, 442
 Chamlee-Wright, Emily 240, 281, 287, 288
 Champ, Patricia A. 399, 403, 421, 534, 539
 Chandra, Amitabh 575, 578
 Chang, Ha-Joon 229, 245
 Chang, Howard F. 529, 539
 Chang, Ruth 524, 539
 charitable donations 342, 448
 charity 13, 22, 44, 60, 61, 62, 63, 90, 143, 144,
 280, 365, 456, 457, 512, 516
 Charness, Gary 195n4, 203, 204
 chartalism 360
 Charusheela, S. 251, 265
 cheating 88, 137n17, 153, 187, 192, 194, 200,
 349, 368
 Cherrier, Beatrice 37, 50
 Chetty, Raj 37, 51
 Chiapello, Eve 338, 356
 childrearing 64
 China 28, 79, 175, 478
 Choi, Seung 29, 32
 Chorus, Silke 260, 266
 Christianity 18, 144
 churning 368
 civic cooperation 193
 civil liberties 327
 civil rights 7, 190, 317, 327
 employment and race 559–78
 civil rights laws
 activities covered 564–6
 policies for future 576–7
 purpose of 560–6
 what do they do? 574–6
 Civil Rights Act 1964 560
 Claassen, Rutger 97, 106n5, 112
 Clark, Gregory 309n14, 313
 Clarke, James J. 368, 376
 climate 49, 108, 143, 150, 299, 303, 399, 436,
 441, 460, 461, 503
 CLIO-Infra project 105
 Coase, Ronald H. 69, 73, 92, 93, 395, 479n4
 Coase Theorem 69
 Coats, A.W. 277, 287
 Code of Professional Conduct (AEA) 87
 coercive paternalism 308
 coercion vs. consent 490–2
 cognition 142
 cognitive science 167
 Cohen, Gerald A. 96, 107, 112, 311n16, 313, 338,
 355, 356
 Cohen, Morrie 355–6
 Cohn, Alain 196, 204
 coinage 360
 Coker, Edward W. 14n5, 30
 Colander, David 155, 156, 229, 245
 Cold War 332
 Coleman, Jules L. 67n26, 73
 collateral 317, 372
 collective decision making 239, 317, 332, 334
 collective ownership 230, 231
 Colman, Andrew xv
 colonialism 301
 Comim, Flavio 401, 421
 commerce 2, 5, 6, 17, 27, 29, 34, 70, 81, 86,
 280, 487
 humanity of 312
 and markets 293–395
 commitment 1, 13n2, 45, 59, 63, 64, 77, 98, 99,
 149, 150, 233, 236, 278, 302, 342, 369,
 395, 487, 533
 commodification 111, 260, 261, 305, 332,
 334, 536
 commodity theory of money 359, 360
 common good 36, 42, 124, 161, 163, 172, 173,
 177, 326, 510

- Commons, John R. 491
 communism 316, 322
 community 19, 20, 27, 38, 57, 59, 79, 86, 87, 89,
 96, 118, 127, 130, 143, 155, 202, 210,
 249, 251, 258, 259, 276, 298, 306, 322,
 338, 339, 342, 483, 484, 485, 502, 509,
 512, 513, 514, 524, 544, 545, 546, 550,
 551, 556
 conflict as foundation of 552–4
 compassion 6, 279, 476, 499, 500, 511, 513, 515,
 516, 517
 fairness, and solidarity 507–10
 compensable vs. noncompensable harm
 481–3
 competition 5, 25, 46, 70, 71, 150–, 163, 172,
 175, 182, 201, 202, 231, 254, 257, 264,
 278, 299, 321, 323, 327, 332, 338, 339,
 343, 344, 346–9, 351–4, 364, 504,
 514, 545
 among suppliers 347
 competitive incentives 195, 195n4
 competitive price-cutting 348
 competitive strategies 354
 complex adaptive system, morality as 138–55
 complex systems
 agent-based adaptation 150–4
 analysis 139
 approach to ethics and economics
 208–27
 two visions 211–13
 macro adaptations 148–50
 complexity 77, 91, 92, 139, 142–8, 277, 283,
 284, 375, 492
 catastrophe 148
 and emergence 146–7
 imperative of 155
 compliance 25, 146, 155, 186, 196, 202, 353,
 531, 536
 behavior, diversity of 143–5
 problems 352
 conceptual interconnectivity 141
 conflict 7, 56, 58, 61, 61n14, 77, 83, 84, 87, 90,
 131, 148, 150, 212, 213, 235, 242, 243,
 244, 277–9, 286, 332, 367, 382, 386,
 437, 464, 471, 483, 511, 515, 516, 529,
 536, 545–8, 562, 563, 573, 597, 598
 environmental 554–6
 as foundation of community 552–4
 of interest, avoiding 368–9
 conformist transmission 127
 conformity 121, 129, 166, 176, 181, 182
 bias 151
 classical emphasis on 164–5
 social 4
 Confucianism 79, 280
 Confucius 89
 congestion 387, 554
 Congleton, Roger D. 333, 335
 Conly, Sarah 307n9, 308, 313
 connectionist theory of the mind 140–1
 Connell, Raewyn 250, 253, 268
 Conrads, Julian 190, 204
 conscience-accounting 186
 consent 67, 232, 233, 236, 240, 243, 299n4, 387,
 456, 477, 523
 vs. coercion 486, 490–2
 consequences 34, 39, 55, 71, 77, 80, 82, 92, 143,
 146, 149, 154, 214, 232, 236–8, 244,
 252, 283, 285, 311, 312, 329, 333, 343n1,
 349, 351, 352, 382, 388, 428–31, 439,
 441, 471, 478, 486, 489, 490, 500–2,
 508, 509, 513, 526, 545, 549
 enriched description of 428
 consequentialism 54, 56, 82, 92, 349, 399, 418,
 522, 524, 525, 525, 527–32, 535
 constraints 56, 59, 60–2, 78, 81, 91, 100–3,
 106–9, 164, 165, 171, 172, 174, 176,
 222, 239, 252n5, 278, 302, 337, 349,
 350, 352–5, 359, 409, 420, 432, 447,
 490, 492, 511, 512, 521, 546
 deontological 522, 524–32, 535, 536
 and thresholds 535–8
 consumers 54, 63, 64, 71, 72, 80, 82, 83, 85, 86,
 118, 201, 202, 258, 325, 333, 344–8,
 355, 367, 368, 369, 375, 383, 391, 431,
 491, 503, 505, 553, 555, 585
 consumption attributes 402, 411n16
 continuity 417–19
 contours of equal goodness 601
 contract theory 182
 contractors 385, 386, 388–91, 395, 514
 contractualism 90
 conversion factors 98, 99, 101, 105, 108, 109
 Cooke, Lynn Prince 258, 266

- cooperation 3, 17, 51, 81, 84, 118–20, 125, 127,
 139, 145–7, 149, 151–4, 174, 176n11,
 193, 202, 302, 305, 549–50, 555, 556
 flexible large group 162–3, 173
 large group 161–4, 173, 176n11
 social 4, 18, 152, 163, 229, 231–3, 236
 in Austrian school 240–4
 Cooter, Robert 521, 539
 Cordes, Joseph 277, 287
 corporate ethics 255, 257, 258
 corporations 338, 339, 341, 342, 343
 corrective economic ethics 254
 corrective-integrative approach to ethics of
 care and ethics of justice 251
 corruption 21, 29, 99, 186, 187, 198, 306, 307,
 310, 333, 352, 448
 potential for 23–6, 28
 cost of lying 187–9
 cost-benefit analysis (CBA) 6, 54, 66, 165, 167,
 168, 170, 173, 176, 213, 217–18, 440,
 478, 478n3, 479n4, 521
 defenses of 410–17
 everyone benefits from in long run 413–14
 rough proxy for overall well-being 414–15
 and social welfare functions (SWF) 399–420
 as SWF 415–16
 and tax system 416–17
 theory 402–4
 Coulhon, T. 597, 605
 courts 333, 342, 545, 552, 563, 565
 Coward, Fiona 163n6, 177
 Cowell, Frank 399, 421
 Cowen, Tyler 41, 42, 51, 240, 244, 245
 Cowton, Christopher J. 374, 376
 Coyne, Christopher J. 2, 7
 Cragg, Wesley 373, 379
 Craig, Lauren 493
 Craswell, Richard 529, 539
 creative destruction 323
 creativity 85, 172, 483, 484
 credit theory of money 360, 362
 creditworthiness 360
 crime 2, 121, 146, 329, 349, 370, 561, 577
 economics of 65
 criminal identity 196
 critical household economics 259
 Crocker, David A. 102, 112
 cross-country differences
 in honesty 197–8
 in lying 191
 cross-disciplinary work 209
 CSR (corporate sustainable
 responsibility) 257
 cultural beliefs 164
 cultural capital 284
 cultural diversity 123
 cultural economics 5, 270, 276
 cultural evolution 124, 149
 cultural group selection 124, 148
 cultural identity 281
 cultural practices 152, 166, 168–71, 278, 283–6
 evolution of ethics from 162–4
 how economic practices may affect 283–4
 influencing of causing economic
 practices 279–83
 cultural sense-making 286
 cultural sphere 272–4
 cultural transmission 129, 130
 culture 5, 20n12, 42, 87, 119, 122–5, 122n6, 142,
 149, 152, 162–4, 196, 238, 259, 298,
 305, 308, 331, 368, 549
 and economy 270–87
 and genes 124, 127–31
 notion of 274–5
 Cummins, Denise Dellarosa 153, 156
 Cureton, Adam 59n11, 73
 Cynics 48
- ## D
- D'Agostino, Fred 139, 155, 156
 d'Aspremont, Claude 404, 421, 422, 426,
 444, 593, 597, 605
 Dahmer, Jeff 563
 Dai, Zhixin 186, 204
 Dallas, Lynne 366, 376
 Daly, Herman 90, 93, 263, 266
 Danby, Colin 259, 266
 Daniels, Norman 507, 511, 518
 Dao, Dr. David 84
 Darley, John M. 531, 542
 Darwall, Stephen 508, 518, 524, 525, 528, 539
 Darwin, Charles 4, 117, 118, 119, 121, 123–7,
 128, 133
 Dasgupta, Partha 35, 51

- Davey, Calum 448, 472
- Davis, John B. xvi–xvii, 2, 5, 7, 208–27
- Davis, Nancy (Ann) 529, 539
- Davis, Scott 93
- Dawkins, Richard 124, 125–7, 126n12, 133, 162n3, 177
- De Waal, Frans B.M. 124, 126, 126n13, 127, 128, 128n15, 133
- Deacon, Terrence W. 129n16, 133
- deadweight loss 85, 333, 345, 354
- Deaton, Angus 585, 605
- Debreu, Gérard 423, 442, 585, 590, 605
- debt 44, 45, 316–18, 332, 339, 361, 362, 366, 367
vs. equity 340
- deception 24, 89, 305, 306, 569
and fraud 367–9
- decision-making 4, 7, 58, 60–3, 65, 66, 72, 150, 160, 162–5, 167, 169, 171–6, 277, 308, 317, 338, 349, 394, 437, 440, 441, 446, 447, 467, 468, 471, 490, 532, 533, 565
- decommodification 261
- degendering economic terms 251–2
- Degrazia, David 523, 539
- dehumanization 304–6
- Dekker, Erwin 233, 245, 280, 287
- delegation 183, 191, 565
and punishment 193
- DeMartino, George F. xvii, 3, 6, 7, 67n27, 87, 88, 93, 227, 474–93, 524n1
- dementia 249
- democracy 5, 68, 85, 236, 238, 392, 393, 395
and capitalism 316–34
caring 251
- democratic capitalism, crises of 332
- democratic self-control (Schumpeter) 324
- democratic states 85, 381
- democratic workplaces 392, 394
- demonetarization 261
- Den Uyl, Douglas J. 15, 18, 19, 30
- deontic weakening 353
- deontology 4, 6, 7, 56, 64, 72, 87, 350, 418, 521, 522, 524–30, 532n4, 536
concerns, indirect ways of
addressing 531–2
constraints and options 522, 524–32, 535, 536
criticism of 528–30
and economic analysis of law 521–38
and Kantian ethics 55–9
moderate 535
absolutist and 526–7, 528–30
and law 530–1, 531n2
threshold 7, 61, 61n15, 527
- deregulation 331n13, 332
- Deschamps, Robert 593, 605
- desire-fulfillment theories of well-being 104
- development ethics 98
- Dhami, Sanjit xv, xvii, 2, 5, 7, 181–203, 204
- Dhawan, Nikita 259, 266
- Diamond, Jared 127, 133
- Diamond, Larry 316, 331, 335
- Diamond, Peter A. 424, 428, 442, 534, 539, 597, 598, 605
- Dichter, Thomas 373, 376
- Dickens, Charles 4, 51
Scrooge and Cratchit [in *A Christmas Carol*] 40–2, 44, 45, 46, 50
- dictator games 197
- dictator(ship)s 85, 190, 193, 194, 197, 200, 393, 407, 407n12, 548
- die-in-the-cup paradigm
lying behavior in 184–5
results of baseline treatment in die-cup experiment 185
- dignity 1, 29, 49, 55–8, 62, 71, 72, 84, 104, 296, 302, 306, 313, 499, 524, 529, 536, 560
and capabilities and moral values 221–4
of labor 389
- Dilworth, John B. 368, 376
- DiMaggio, Paul 279, 281, 282, 287
- Dionne, G. 441
- direct normative prescription 352–3
vs. adversarialism 353
- disapproval 164, 166, 168, 170, 171
- discourse ethics 249
- discrimination 2, 7, 64, 196, 249, 484, 503, 508, 509, 515, 516, 523, 524, 531, 532, 534–7, 560, 561–3, 568–77
disparate treatment discrimination 568–9
results 569–71
disparate impact discrimination 568
results 571–4
- dishonesty 14, 89, 186, 187, 192, 194, 198, 349n2
slippery slope of 197

distribution 29, 82, 107, 110, 184, 186, 187, 193,
198, 234, 237, 252, 253, 258–60, 264,
265, 276, 323, 354, 365, 389, 399, 402,
414, 416, 428–34, 438, 440, 446,
450–68, 470, 471, 486, 493, 500,
513, 515, 516, 524, 546, 559, 559n1,
587, 91–6
of harm 489
interpersonal prospect 452–3
interstate 452, 453–4
intrapersonal prospect 452, 454
intrastate 452, 454
mechanism of harm generation
and 489–90
distributional sensitivities 446, 447, 450–4,
455, 459, 468, 470
prospect values, and parameters 454
distributive costs 333
distributive justice 96, 98, 99, 143, 220, 233,
425, 551
disutility 183, 199–201
diversification 447–8, 460
and sensitivity to interstate
distribution 456–7
Dobb, Maurice 325
Dodge v. Ford 341
Dolan, Paul 402, 421, 502n3, 503n4, 518
Doleac, Jennifer 575, 578
Dolfsma, Wilfred 2, 7
Dolinko, David 528, 539
domestic labor debates 253
Donaldson, David 452n3, 472, 604
Donath, Susan 252, 264, 266
Donohue, John J. III 531, 532, 534, 539, 540
Dorff, Michael B. 523, 531, 540
Dorsey, Dale 601, 605
double coincidence of wants 359
Douglas, Mary 282
Dowding, Keith 110, 112
Dowling, Grahame R. 14, 32
dual systems theory 253
Dufwenberg, Martin 191, 199, 200, 203, 204
Dunbar, R.I.M. 163n6, 177
Durham, William H. 124, 133
Durkheim, Emile 284, 287
Durlauf, Steven N. 564, 578
Durnin, Maureen 440, 442

Duska, Ronald F. 368, 369, 376, 379
Dutch Republic 280
Dutt, Amitava 3, 7
duty 4, 57–61, 61n14, 64, 65, 68, 77–9, 81, 87, 88,
90, 121, 173–7, 324, 364, 365, 368–72,
374, 478, 480, 481, 485, 528, 552
moral restraint, based on 173–7, 176n10
imperfect 57, 63
Kantian 55, 60, 79, 84, 525
perfect 57, 60, 63
Dworkin, Ronald M. 68, 73, 425, 426, 438,
442, 536, 540
Dwyer, John 14n5, 30
dynamic efficiency 86
dynamic justice 551–2

E

Easterbrook, Frank H. 342, 356
Ebola 460
Eckel, Catherine C. 196, 204
ecological justice 551, 557
ethics and economics of 544–57
econogenic harm 6, 475, 492
sources of 476–8
econometrics 91
economic analysis 1, 5, 14, 29, 35–7, 49, 50, 65,
69, 138, 216, 223, 229, 230, 234, 252,
277, 281, 476, 500, 510, 517, 554
of law
and deontological morality 521–38
integrating thresholds, constraints, and
options with 535–8
economic assessment, harm-centric
approach to 485–92
economic behavior 2, 4, 5, 40–2, 64, 200, 209,
223, 270, 281, 286
effect of ethics on 172–4
effect of group size on 170–2
evolution of 164–70
and ethics and rationality 160–73
economic development 282, 285, 286, 333
economic efficiency 78, 86, 92, 390, 521
economic ethics 5, 248, 255–6, 258, 259,
262, 264, 266
and capability approach 96–111
corrective 254
feminist *see* feminist economic ethics

- economic ethics (*cont.*)
 functionalist 254
 integrative 254
 overcoming androcentric 254–5
- economic goals and methodologies for
 human well-being 85–7
- economic harm
 complex nature of 483–5
 elements of profile 486
 nature of and responsibilities of
 economists 474–93
 standard approach 478–9
- economic liberalism 329
- economic man 117, 120, 131–2
- economic outcomes 41, 259, 285
- economic practices 278, 287
 cultural practices influencing or
 causing 279–83
 how they may affect cultural
 practices 283–4
- economic prosperity 297
- economic research 87, 277
 development as akin to religion 89
- economic systems 4, 103, 105–11, 139, 256,
 263, 264, 316, 476, 487–90, 492
 alternative 320, 333
- economic terms 65, 276
 degendering 251–2
- economic theory 2–4, 65, 117, 259, 263, 337,
 338, 368, 423, 584, 585, 587
 dangers of ignoring 596–599
 of paid and unpaid work 261–2
- economic way of thinking 40–6, 237, 277
 Dickens's *Scrooge* example 40–2
- economics
 cultural 5, 270, 276
 as cultural practice 275–8
 as engineering 239
 and ethics 1–7
 Adam Smith and study of ethics in
 commercial society 13–29
 capability approach 96–111
 complex systems approach 208–27
 of ecological justice 544–57
 economic ethics and capability
 approach 96–111
 ethical pluralism in economics 77–93
 evolution and moral motivation in
 economics 117–32
 on evolution of ethics, rationality, and
 economic behavior 160–77
 feminist 248–65
 and health insurance 499–517
 Kantian approach to economics 54–73
 mainstream and social economic
 visions of 213
 morality as complex adaptive
 system 138–55
 rethinking Hayek's social ethics 138–55
 as single subject of investigation 208–11
 future prospects for 226–7
 virtue and economics 34–50
 is it value free? 232–7
 lessons from 583–604
oikonomia 251
 performative 276
 as sense-making activity 276
 social 5, 210, 211, 213,
 understanding of positive-normal
 distribution 219–21
 tragedy of 474–93
 undoing gender in 251–5
 and virtue 34–8
 without economic man 131–2
- Economics Job Market Rumors website 87
- économie sociale* 210
- economists
 as ethicists 87–9
 code of Professional Conduct (AEA) 87–8
 why a code is not sufficient 88–9
 questions and answers about labor
 markets 567–76
 responsibilities of, and nature of economic
 harm 474–93
 what do they know about race in labor
 market 566–76
- Econs 181–5, 188, 191, 193, 195, 201
 as amoral and self-regarding 181
- Edelman, Lauren B. 575, 578
- Edgeworth, Francis Y. 13n1, 93, 131, 133
- Edmonds, David 56n3, 73
- education 5, 26, 86, 97, 103, 107, 131, 191, 221–3,
 226, 272, 284, 298, 323n6, 373, 499,
 501, 503, 562, 565, 566, 570, 572–4, 577

- effective freedom 101–3, 107, 110, 111, 491
 efficiency 6, 45, 78, 79, 82–6, 90, 92, 163, 208,
 213, 215–18, 221, 225, 252n5, 262, 276,
 278, 333, 333n14, 354, 355, 390, 392,
 393, 395, 423, 426, 485, 492, 500, 501,
 503, 511–17, 521–3
 dynamic and static 86
 ethical limits on 507–10
 in health care 500–3
 Kaldor-Hicks 66, 67n27, 68, 70, 411–13, 478
 as “normative” concept 215–16
 egalitarianism 6, 96, 311, 425, 428, 429, 434,
 435, 438, 440, 441, 589
 Eggleston, Ben 418, 421
 elections 190, 191
 electoral politics 333
 Elias, Julio Jorge 86, 93, 275, 287
 Elgin, Catherine 220n4, 227
 Ellerman, David 392, 396
 Ellis, Anthony 530, 540
 Ellsberg, Daniel 437, 442
 Elson, Diane 249n2, 258, 261, 262, 266
 Elster, Jon 130, 133, 318, 321, 335, 483, 488, 494
 Emami, Zohreh 227
 embedding of economy in society 328
 emergence and complexity 146–7
 emergent property 146–8
 emotional contagion 127, 127n14
 emotions 15, 19, 121, 123, 126–8, 130, 169, 183,
 197–200, 301, 437
 empathy 127, 129, 176
 problem 162n4, 170
 employment 16, 25, 26, 67, 88, 217, 237, 252,
 273, 330n9, 355, 383–6, 388, 389n5,
 390–2, 395, 485, 531, 537
 civil rights and race 559–78
 discrimination: how does it violate
 someone’s rights? 561–3
 Enders, Walter 88, 93
 Engel, Antke 266
 Engelen, Peter-Jan 370, 377
 Engelhardt, H. Tristram, Jr. 500n2, 511, 518
 Engels, Friedrich 284, 287
 Engstrom, Stephen 59n11, 73
 entanglement thesis 220, 225
 entrepreneur-driven innovation 323
 entrepreneurship 273, 280, 281, 283, 285
 environmental conflict 554–6
 environmental damage 109
 environmental economics 86
 environmental sustainability 374
 Epictetus 39
 Epstein, Larry G. 428, 442
 Epstein, Richard A. 70, 73, 524, 535, 540
 equal goodness, contours of 601
 equality 3, 6, 17, 56, 66, 97, 106n4, 141, 184,
 222–4, 250, 256, 258, 302, 304,
 308–10, 318–20, 370, 418, 419, 424,
 429–34, 438, 466, 491, 509, 524, 552,
 584, 589–93, 595–7, 599, 604
 of conditions 318
 of status 318
 of what? 96
 equilibrium 35n1, 64, 71n30, 117, 151n14,
 162n3, 191, 202, 225, 226, 236, 323,
 346, 348, 351, 382, 384, 423, 440, 585
 equity 208, 257, 342, 343, 361, 395, 427, 428,
 435, 439, 440, 464, 465
 capital 340
 vs. catastrophe avoidance 432–3
 vs. debt 339, 340
 Erat, Sanjiv 183, 187, 188, 193, 204
 Esping-Andersen, Gøsta 321, 335
 Estevez-Abe, Margarita 321, 335
 ethical behavior 2, 3, 5, 37, 38, 40, 58, 63, 64,
 81, 174, 192, 193, 202, 212
 fitting virtue to 46–9
 ethical concerns 14, 37, 231, 234
 and economic analysis 237–40
 ethical dissonance 88, 198
 ethical impartiality 417
 ethical limits on efficiency 507–10
 ethical pluralism in economics 77–93
 ethical theories in economics 13–111
 ethical turn in feminist theory 251
 ethicality, human 181–203
 ethically benign vs. ethically indictable
 harms 488
 ethics
 and Adam Smith
 in commercial society 13–33
 ethics and markets 20–8
 limits of benevolence 21–3
 possibility of moral development 26–8

ethics (*cont.*)

- potential for corruption 23–6
- theory of 18–20
- in applied economics 497–578
- bridging ethical boundaries 78–81
- of care 251
- deontological 4, 6, 7, 56, 64, 72, 87, 350, 418, 521, 522, 524–30, 532n4, 536
- and economics 1–7
 - Adam Smith and study of ethics in commercial society 13–29
 - capability approach 96–111
 - complex systems approach 208–27
 - of ecological justice 544–57
 - economic ethics and capability approach 96–111
 - ethical pluralism in economics 77–93
 - evolution and moral motivation in economics 117–32
 - on evolution of ethics, rationality, and economic behavior 160–77
 - feminist 248–65
 - and health insurance 499–517
 - Kantian approach to economics 54–73
 - mainstream and social economic visions of 213
 - morality as complex adaptive system 138–55
 - rethinking Hayek's social ethics 138–55
 - as single subject of investigation 208–11
 - future prospects for 226–7
 - virtue and economics 34–50
 - within Austrian school of economics 229–45
- effect on economic behavior 172–4
- ethical pluralism in economics 77–93
- of ethics from cultural practices 162–4
- feminist 248–51
- financial 358, 359, 375
- fitting virtue to ethical behavior 46–9
- of justice 251
- and, in, for labor markets 381–95
- of making risky decisions for others 446–71
- of money and finance 358–76
- and rationality and economic behavior, evolution of 160–73

- and science 220
- in schools of economics 179–287
- of technology 97
- towards real world ethics 556–7
- utilitarian 54, 59, 66–8, 72, 78, 80, 82, 90, 91, 123n8, 149, 418–20, 424–7, 430, 434, 435, 438–40, 446, 508, 523, 535, 569, 585, 595–7, 599, 601
- virtue ethics 18, 35, 37–41, 42n5, 44–7, 59, 59n11, 78, 79, 81, 87, 209, 210, 280, 532n4
- Etzioni, Amitai 2, 8, 62, 63, 63n18, 73, 74, 119n5, 133
- eudaimonia* 79
- Europe 17n7, 28, 210, 233, 252, 275, 279, 280, 316, 394
- Eusepi, Giuseppe 535n5, 539
- Evans, Nicholas 129n16, 133
- Evensky, Jerry 123, 133
- evolution 4, 7, 139, 139n1, 142, 148–51, 153, 154, 323, 553n9
 - of ethics
 - from cultural practices 162–4
 - and rationality, and economic behavior 160–77
 - and moral motivation in economics 117–32
- evolutionary selection 126n12, 148, 546
- evolutionary studies 118
- evolved complex system, reverse engineering of 150
- ex ante egalitarianism 425, 428–9, 435, 440
- ex post egalitarianism 425, 429–32, 435, 440, 441
 - vs. separability 434
- ex post inequalities 424, 440
- exchange 14, 16, 27, 35n1, 40, 64, 69, 71, 82, 141, 172, 202, 210, 220, 233, 236, 237, 240, 243, 273, 281, 286, 295, 296, 298, 300, 301, 303, 305, 307, 308, 321, 332, 345, 359–61, 363, 365, 374, 375, 381, 382, 490, 491, 553, 555
- exchange value of labor 320, 321
- existential ethics 250
- experimental economics 81, 88, 126, 176, 183
- exploitation 23–5, 25n18, 29, 111, 163, 253, 264, 320, 321, 338, 365, 484, 510

externalities 88, 201, 202, 213, 216–18, 235,
239, 308, 387, 388, 433, 521, 555
rights and wrongs 67–72
extrinsic motivation 200
extrinsic preferences 182

F

Fabozzi, Frank J. 361, 377
facial expression 127
fair allocation theory 423
Fairlie, Robert W. 573, 578
fairness 6, 37, 66, 78, 79, 83, 86, 120, 123, 126,
128, 130n17, 139, 143, 151, 164, 181, 217,
308n11, 392, 395, 423, 427, 430,
439–41, 486, 500, 511, 513, 515–17,
523, 524., 527, 533, 561, 595, 597
in financial markets 367–70
solidarity and compassion 507–10
traffic management before 551–2
vs. utilitarianism 424
Falk, Armin 201, 204
family 16, 59, 61, 64, 65, 77, 92, 105, 125, 143,
166, 185, 249n2, 252, 257, 258, 260,
271, 273, 282, 284, 286, 318, 319, 323,
338, 369, 400, 405, 418, 420, 433,
500, 509, 524, 562, 574, 574n6
Faravelli, Marco 195n4, 204
fascism 139n1, 328, 329
Faulhaber, Gerard R. 276, 287
fear 21, 127, 170, 175, 257, 319, 484, 576
feedback loops 146
Fehr, Ernst 2, 8, 36, 51, 118n3, 120, 126n13,
128n15, 131, 133, 182, 191, 193, 197,
204, 205
Feinberg, Joel 481, 485, 494
fellow-feeling 15, 18, 301
feminist critical reflection 255–9
feminist economic ethics 248–65
method, subject matter, and model of
action 255–64
methodological approach 255–9
individual level 258–9
organizational level 257–8
structural level 256–7
reflected subject matter: paid and
unpaid work 259–62
feminist economics 5

of caring provisioning 264
and ethics 248–65
feminist institutionalist economics 253
feminist Marxist economics 253
feminist neoclassical economics 253
normative foundations of 252–4
feminist postcolonialist ethics 251
feminist theory, ethical turn in 251
Ferber, Marianne A. 2, 8, 253, 266
Ferguson, Niall 360, 377
Ferranna, Maddalena 433, 435, 440, 441, 442
Ferzan, Kimberly Kessler 531, 540
Fessler, Daniel M.T. 197, 205
Feynman, Richard 182
fictitious commodities 328, 329
fiduciary duty 80, 368–70
Field, Alexander J. 118n2, 128n15, 133
finance
ethics of money and 358–76
social responsibility of 371–5
what is it? 361–3
financial assets 361–2
financial crises 362, 375, 487
and systemic risk 371–2
financial ethics 358, 359, 375
financial markets 92, 282, 283, 359, 362,
374, 375
fairness in 367–70
financial risk 367, 368, 371, 372, 374, 504
financial services 307, 372, 373
Fine, Ben 64, 74
Finkelstein, Amy 425, 442, 504, 518
Fischbacher, Urs 120, 131, 133, 184, 186, 187,
189, 193, 200, 203, 205, 207
Fischel, Daniel R. 342, 356
Fischer, Liliann 3, 8
Fishburn, Peter C. 593, 605
Fiske, Alan Page 128, 134
five spheres model 271–4, 272
fixed population framework 400
Fleischacker, Samuel 14n5, 20, 23n15, 24, 26, 30
Fletcher, Guy 401, 421
Fleurbaey, Marc xvii, 6, 66n24, 308n12, 313,
415n20, 420, 421, 423–41, 442, 443,
445, 452n3, 458n4, 461n7, 466n8,
471, 472, 593, 605
Flew, Anthony 338, 343n1, 356, 364, 377

- Florida, Richard 270, 283, 287
 Folbre, Nancy 77, 94, 104, 112, 263, 266
 Follmi-Heusi, Franziska 184, 187, 189, 200, 205
 Foot, Philippa 55, 74, 526, 540
 Force, Pierre 23, 30
 Ford Motor Corporation 341
 foreseeable vs. unforeseeable harms 487
 formalism 231
 Forman-Barzilai, Fonna 20n12, 21, 22, 30
 Formula of Respect (Kant) 56, 57, 58, 67
 Formula of the Kingdom of Ends (Kant) 56
 Formula of the Law of Nature (Kant) 57n7
 Formula of Universal Law (Kant) 56, 57, 58
 Fortenbaugh, William 40, 47, 51
 fortune 301, 312
 influence of (Smith) 311, 311n18
 instability of (Hume) 311
 Foucault, Michel 277, 278, 287
 "Four Horsemen" of leveling (Scheidel) 310
 four sector model, shifting processes of
 work in 261
 four spheres (Polanyi) 271
 France 143, 299, 318, 503
 Frank, Robert H. 81, 94, 168, 169, 177, 356
 Frankfurt, Harry G. 308n13, 313
 Frankfurt School 283
 Fraser, Clive xv
 Fraser, Nancy 260, 266
 fraud 71, 72, 198, 296, 301, 302, 309, 369
 and deception 367–8
 free institutions 319
 free-market capitalism 328
 free-market systems 161
 free speech 69, 85, 394
 free trade 108, 299, 329
 freedom 4, 45, 48, 61, 79, 84, 86, 97, 98–111,
 98n1, 220–4, 231, 236, 240, 242, 256,
 280, 304, 317, 319, 319n4, 320, 322,
 323, 327–30, 386, 394, 481, 483, 484,
 492, 500, 507, 514, 517, 523, 524, 531,
 536, 545, 560–2, 564, 565, 577, 578
 and capability approach 100–3
 and choice in health care 511–13
 effective 101–3, 107, 110, 111, 491
 of the press 85, 319
 and well-being 97
 Freeman, A. Myrick 402, 421
 Freeman, Samuel 528, 540
 French Revolution 149, 232, 318
 Frey, Bruno S. 36, 51, 81, 94, 120, 132, 134, 270,
 277, 284, 288
 Frick, Johann 461n7, 472
 Fried, Jesse 343, 356
 Friedman, Daniel 142n5, 156
 Friedman, David 90, 94
 Friedman, Milton 80, 81, 273, 491, 494
 Fryer, Roland G., Jr 571, 572, 578
 Fujiwara, Daniel 402, 421
 functionings 98–105, 98n1, 103n3, 108, 110, 256
 achieved 98, 99
 bundles of 98
 Fuster, Joaquin 161, 161n1, 177
 functionalist economic ethics 254
- ## G
- Gächter, Simon 36, 51, 118n3, 128n15, 133, 193,
 198, 204, 205
 Gajdos, Thibault 428, 431, 438, 443
 Galbraith, John Kenneth 348, 356
 Galtung, Johan 488n10, 494
 game theory 139, 151, 161, 183, 276, 585
 psychological 198–200
 Garbarinoa, Ellen 193, 205
 Garber, Marjorie 251, 266
 Gärdenfors, Peter 140, 156
 Gardiner, Coral 29n20, 30
 Garrett, Neil 197, 205
 Gates, Bill 309
 Gaus, Gerald xvii, 4, 29, 30, 55n2, 74, 138–55,
 156–7, 240, 552, 558
 Gauthier, David 446n1, 472
 Gawande, Atul 510n7, 518
 Geertz, Clifford 274, 278, 286, 288
 Geistfeld, Mark A. 70, 74, 480n5, 482, 494
Gemeinwirtschaft 210
 gender 201, 248, 251, 256–8, 260, 262–5,
 565, 569
 differences, in lying 191, 198
 in economics, undoing 251–5
 norms 5, 256
 critical reflection of 249–50
 relations 250
 roles 142
 gendered earner model 260

- generalized utilitarian family of SWFs 400, 418
- generosity 21, 22, 41, 126, 298
- genes 124–31, 161, 162n3, 165, 168, 548, 549
and culture 124, 127–31
“ruthless selfishness” of 126
- genetic endowments 130
- genetic evolution 142
- genetic group selection 124
- George, David 62, 74
- German Historical School 253
- Germany 185, 186, 299, 331, 503
- Gershman, John 373, 377
- Gevers, Louis 404, 421, 593, 605
- Ghalib, Asad 373, 378
- Ghiselin, Michael 126, 132, 134
- Gibbard, Allan 201n2, 30, 469, 472
Libertarian Paradox 469–70, 471
- Gibson, Rajna 187, 205
- Gibson-Graham, J.K. 259, 266, 492, 494
- Gideon, Jasmine 249n2, 258, 262, 266
- gig economy 6, 382, 394
labor markets and rise of 385–91
- Gilbert and Sullivan’s *Trial by Jury* 469
- Gilboa, Itzhak 402, 421, 436, 443
- Gilman, Charlotte Perkins 248n1, 266
- Gil-White, Francisco J. 128, 135
- Gilligan, Carol 250, 266
- Gini function 430, 559n1
- Gino, Francesca 188, 205
- Gintis, Herbert 2, 7, 36, 51, 81, 93, 94,
118nn2,3, 120, 125, 128, 128n15, 131,
131n18, 133, 134, 142n5, 148, 150,
151nn13,14, 156, 182, 183, 190, 191, 193,
203, 205, 308n11, 313, 381, 395
- Giroux, Henry 338, 356
- Giuliano, Paola 270, 278, 282, 285, 286, 287
- globalization 108, 264, 331n13
- Gneezy, Uri 81, 94, 182, 183, 184, 187–90,
191, 197, 201, 204, 205, 206
- Goh, Eugenia 493
- Goldberg, Abbie E. 259, 266
- golden rule 249
- Goldsmith Arthur H. 574, 579
- Gollier, Christian 433, 443
- Goodin, Robert E. 480, 488, 494, 531n2, 540
- Goodman, Nelson 477n2, 494
- Google 64, 307n2, 391
- Görgemanns, Herwig 46, 51
- Gould Ellen, Ingrid xix
- Gould, Skye 86, 94
- governance 273, 275, 394
- government
intervention 72, 329, 333
regulation 307, 333, 338
transfers 333
- governmental decision-making 308
- Gowdy, John 153, 154, 157
- Graafland, Johan J. 365, 369, 377
- Graeber, David 360, 361, 377
- Graham, Carol 401, 421
- Graham, Julie 260, 269
- Granovetter, Mark 14, 28, 30
- Grant, Simon 438, 443
- Grantham, George 549, 558
- Great Depression 236, 358
- Greaves, Hilary 452, 472, 597, 605
- Greece, Ancient 79, 160, 176
- greed 14, 29n19, 241, 297, 318n3, 343, 363,
364, 366, 368, 373, 563
- Green, Sara 150, 157
- Greene, J.D. 126n13, 134
- Greenfield, Kent 338, 356
- Gregor, Mary J. 57n9, 74
- Greif, Avner 281, 288
- Griffin, James 401, 421, 501, 518, 523, 528, 540,
602, 605
- Griswold, Charles L. 15, 18, 19, 22n14, 26, 30,
38n4, 39, 51
- Grolleau, Giles 192, 205
- group
approval 164
competition 150
disapproval 164
-level evolution 153
-level selection 150
rationality 169, 171, 173
size, effect on economic behavior 170–2
- groups, lying in 194–5
- Grüne-Yanoff, Till 60n13, 74
- Gudeman, Stephen 272, 288
- guilt 65, 120, 121, 126, 162, 166, 170, 181, 183,
188–90, 197–201
aversion 188, 197, 198, 200
- Gunning, James Patrick 231, 245

Güth, Werner 118n2, 126n13, 134

Guzmán, Ricardo Andrés 128n15, 134

H

Habermas, Jürgen 249, 271, 272, 273, 288

Haeffele, Stefanie xvii–xviii, 4, 13–29, 25n18, 29, 33

Haidt, Jonathan 122n7, 126n13, 128, 130, 134

Haigh, Matthew 374, 377

Haley, Kevin J. 197, 205

Hall, Joshua C. 308n11, 313

Hall, Peter 275, 288, 331, 335

Hallsworth, Michael 195, 206

Hamilton, William D. 125, 126n12, 134

Hamlin, J. 128, 134

Hammerstein, Peter 118n2, 134

Hammond, Peter J. 424, 426, 443, 593, 605

Hancké, Bob 331, 335

Hands, Wade xvii, 227

Hankins, Keith 311n18, 313

Hanley, Ryan Patrick 21, 30, 39, 51

Hanlon, Niall 251, 266

Hanna, Rema 187, 206

Hansen, Benjamin 575, 578

Hanson, Robin 40, 52

Hansson, Sven Ove 60n13, 74

happiness 20, 22, 23, 48, 54, 66, 82, 90, 91, 105, 164, 176, 218, 243, 253, 301, 313, 326, 337, 355, 401

Harari, Yuval Noah 161, 173, 177

Hardin, Russell 139, 142, 143, 157

Hare, Richard M. 120, 134

Harel, Alon 520, 540

Harford, Tim 41, 51

harm

avoidable vs. unavoidable 476, 486, 487

-centric approach to economic assessment 485–92

compensable vs. noncompensable 481–3

distribution of 489

econogenic 475

sources of 476–8

economic

complex nature of 483–5

elements of profile 486

and nature of and responsibilities of economists 474–93

standard approach 478–9

ethically benign vs. ethically indictable 488

foreseeable vs. unforeseeable 487

iatrogenic 475, 477

mechanism of generation and

distribution 489–90

nature of prevalent, prevented, and covered harms 487

necessary vs. unnecessary 487

productivity of 487

reparable vs. irreparable 487

taxonomy of harmed or harmful

conditions 484

vs. wrongs 67–70

Harper, Malcolm 373, 376

Harris, John 531, 540

Harris, Seth D. 389, 390, 391, 396

Harrison, Lawrence 280, 288

Harsanyi, John C. 423, 425, 426–8, 430, 432, 438, 439, 443, 444, 451, 472, 587–8, 593, 594, 596–7, 598, 605

Harsanyi's theorem 426–8, 434, 435, 451

Harsh, Mel 44, 51

Harter, Jim 396

Hartmann, Heidi 253, 266

Hauser, Marc D. 126n13, 129, 134

Hausman, Daniel M. xviii, xix, xxi, 3, 6, 8, 36, 51, 138, 157, 215, 217, 227, 499–517, 523, 534, 535n5, 540

Hausman, Jerry A. 534, 539

Hayashi, Takashi 428, 442

Haybron, Daniel M. 401, 421

Hayden, Amanda 493

Hayek, Friedrich von 4, 5, 27, 30, 35–6, 51, 106, 112, 138, 139–40, 139n1, 142, 145–50, 152–5, 157, 176n11, 177, 230, 231, 235n2, 236–9, 240, 243, 244, 245, 246, 273, 277, 280, 324–8, 329, 330, 331n11, 335

on capitalism and democracy 324–8

Hazelton, James 374, 377

Hazlitt, Henry 242, 243

Heacock, Marian V. 368, 377

Heal, Geoffrey 436, 444

Heath, Joseph xviii, 5, 337–56

health care 6, 107, 222n5, 394, 499–507, 509–17, 565

efficiency in 500–3

freedom and choice in 511–13

- and health insurance markets,
 fixing 513–17
 spending 86
- health insurance 6, 425, 426, 440,
 economics and ethics 499–517
 markets and health care, fixing 513–17
- Heatherton, Todd F. 62, 73
- Hebb, Donald 167n8, 177
- Heckman, James 569, 575, 579
- hedonistic theories of well-being 104
- Hees, Martin van 110, 112
- Hegel, G.W.F. 488
- Heimbach-Steins, Marianne 249, 266
- Helbing, Dirk 118n3, 131n18, 134
- Hellman, Deborah 536, 540
- Henderson, Monika 29n20, 30
- Hendry, John 364, 366, 368, 377, 379
- Henrich, Joseph 36, 42, 51, 118n2, 124, 128,
 128n15, 134, 135, 142, 142nn2,3,
 149, 152–3, 157
- Herman, Arthur 279, 288
- Herman, Barbara 58, 74
- hermeneutic methods 278
- Herne, Kaisa 439, 444
- Herrmann, Benedikt 193, 206
- Herrnstein, Richard J. 573n5, 579
- Herzog, Don 535, 540
- Herzog, Lisa 14n5, 17, 30, 359n1, 366, 369, 377
- heterogeneity 144, 191, 492
- heteronormativity 251, 259, 262, 264
- heterosexuality 251
- Hicks, John R. 54, 74, 411, 421, 474, 475, 476,
 478, 479, 480, 485, 494
- hierarchical workplaces 393
- High, Jack 234, 246
- Hild, Matthias 426, 444
- Hill, Kim R. 124, 135
- Hill, Lisa 23, 30
- Hill, Thomas E., Jr. 56, 56n5, 59n11, 73,
 74, 558
- Hiller, Janine S. 74n19, 74
- Hirschman, Albert O. 272, 288
- historical analyses 105
- history-dependent lying 194
- HIV 460
- Hobson, John A. 132, 135
- Hodgson, Geoffrey M. xviii, 4, 117–32, 135
- Hodson, Randy 29n20, 30
- Hoffman, Mitchell 308n11, 313
- Hofstede, Geert 270, 281, 282, 288
- Holland *see* Netherlands, the
- Holland, John H. 147, 147n10, 148, 158
- Holmes, Oliver Wendell 65, 74
- Holtug, Nils 418, 422
- home economics 259
- home sphere 272
- homo behavioralis* 182, 183, 202
- homo economicus* 13, 56, 81, 140, 181–3, 188,
 202, 209, 253, 254
- honesty 14, 29, 34, 38, 60, 62, 87–9, 181,
 185, 186, 188, 189, 194, 196, 352,
 367, 369
 cross-country differences in 197–8
- Hont, Istvan 23, 31
- Hooker, Brad 532, 540
- Hoover, Gary A. 88, 93
- Hope, V.M. 18, 31
- Hoppe, Hella 253, 266
- horizontal and vertical pluralism 78–80, 79
- Horkheimer, Max 283–4, 288
- Horner, Victoria 152, 158
- Hosein, Adam O. 531n2, 535, 536, 540
- Hotelling, Harold 411, 422
- household ethics 258
- Houser, Daniel 194, 198, 206
- housewifization 261
- Howard-Snyder, Frances 525, 540
- Hoy, Michael 440, 441, 444
- Hudon, Marek 373, 377
- Hudson, Richard 374, 377
- Hugh-Jones, David 197, 206
- human action 89, 142, 153, 242
- human behavior 3, 35n1, 78, 161, 182,
 522, 535n5
 and modeling markets 80–1
- human cooperation 119, 152, 244
- Human Development Index (HDI) 86
- human ethicality 181–203
- human material of politics (Schumpeter) 324
- human morality 124, 182, 200, 202
 exploring richness of 191–8
- human rights 49, 72, 78, 79, 85, 222–4, 249,
 373, 524, 560
- human sociality 200, 202
- human welfare 78, 523, 528, 534, 538
 consequential evaluations of 82–7

human well-being
 economic goals and methodologies for
 85–7
 humane markets 295–313
 can markets be humane? 310–13
 humane political economy 296–303
 of Hume 297–9
 of Smith 299–303
 Hume, David 5, 18, 18n8, 24n17, 139, 151,
 213–14, 214n2, 219, 220, 227, 228,
 232, 240, 296, 297–9, 299n4, 303,
 304, 305, 306, 309–13, 314
 humane political economy of 297–9
 Humean guillotine 214
 Huntington, Samuel P. 275, 280, 288
 Hurka, Thomas 401, 422
 Hussain, Waheed 350, 351n3, 356, 357
 Hutcheson, Francis 18, 24n17
 Huxley, Julian 126, 135
 Huxley, Thomas Henry 125–7, 126n11, 135
 hypocrisy 296, 287, 306

I

Iacobucci, Edward M. 531, 540
 iatrogenic harm 475, 477
 Ibrahim, Solava 100, 112
 ICESCR (The International Covenant on
 Economic, Social and Cultural
 Rights) 249, 249n2
 identity 48, 64, 77, 78, 132, 175, 275, 400,
 451, 562
 criminal 196
 cultural 281
 professional 196
 social 183, 189, 191, 196
 and morality 196
 Ignatieff, Michael 23, 31
 ignorance 122, 193, 233, 237, 238, 437, 477,
 478, 492
 veil of 425, 426, 438–9, 440, 508, 529
 imitation 130, 152
 conformist transmission 127
 prestige-based 128
 impartial spectator 15, 19–21, 166, 249
 imperative of complexity 155
 imperfect duty 57, 63
 imperialism 301

incentive compatible elicitation 184
 incentives 13, 35, 62, 80, 88, 110, 119, 182–4,
 191, 192, 195, 195n4, 225, 235, 278,
 343, 346, 352, 353, 450n2, 514, 521,
 531, 563, 576
 and lying 187
 and markets and morality 200–2
 independent contractors 390, 391, 395, 514
 independent workers 390
 India 79, 187, 372, 395
 indifferents (Stoic) 47–9, 48n7
 individual rationality 151, 163, 167, 169, 171,
 173, 176
 individualism 81, 87, 121, 130, 160, 174–6, 254,
 319, 320, 322
 Industrial Revolution 29, 259, 328, 329
 inequality 3, 37, 99, 130, 234, 283, 304, 308–10,
 334, 365, 373, 399, 425, 427–30, 432,
 433, 435, 439–41, 451, 456, 476, 480,
 489, 509, 554, 555, 559n1, 583, 592
 influence of fortune (Smith) 311
 inheritance (genetic) 124, 127
 injury 24, 72, 308, 449, 450, 483–5, 499, 517
 injustice 3, 23, 237, 250, 254, 296, 297, 309, 313,
 485, 552n7, 560
 Innes, Mitchell 360, 377
 innovation 3, 17, 24, 26, 85, 103, 125, 275, 280,
 323, 354, 58, 476, 503
 insider trading 367, 369–70
 instability of fortune (Hume) 311
 integrative economic ethics 254–6
 interdisciplinarity 139, 209
 interdisciplinary analysis 138, 139–40
 interest rates 277, 366, 373
 intergroup competition 163, 172
 intergroup migration 124–5
 interpersonal prospect distribution 452–3
 sensitivity to 455–6
 critics against 463–4
 and restrictive intervention 455–6
 interpretive methods 278, 281
 interstate distribution 452, 453–4
 sensitivity to 456–7
 critics against 464–5
 and diversification 456–7
 intervention 6, 71, 72, 86, 149, 219, 221, 226,
 230, 235, 329, 333, 446, 448, 476, 477,

477n2, 478, 480, 482, 483, 483n8,
490, 492, 503, 510, 534, 559
restrictive 447, 455–6, 460, 462
interventions-based social policies 225–6
intrapersonal prospect distribution 452
critics against 466–8
sensitivity to and not on my watch 458–9
intrapersonal well-being comparisons 404
intrastate distribution 452, 454
sensitivity to 457–8
critics against 465–6
and unwelcome risk reductions 457–8
intrinsic motivation 37, 120, 200, 201
intrinsic value 102, 328, 362, 521, 523, 524
investment 230, 278, 283, 321, 325, 340, 341,
353, 366, 367, 368, 371, 383, 461,
487, 531
socially responsible 373–5
investors 258, 340–3, 361, 362, 366, 374
involuntary unemployment 389
and labor markets and legitimate
wages 382–5
Inwood, Brad 43, 51
Ireland, Thomas R. 482n7, 494
Irigaray, Luce 250, 267
irregularity of sentiments (Smith) 312
irreparable ignorance 477, 478, 492
Irvine, William B. 369, 374, 377
Irwin, Terence 56n5, 74
is-ought doctrine 214, 219
“is” type statements 214, 219, 220
Islam 279
Ismael, Jenann 155
Italy 299
Iversen, Torben 330n10, 331, 335

J

Jacobsen, Joyce 259, 267
James, Aaron 371, 372, 377
Japan 28, 280
Jaworski, Peter M. 307n9, 308n11, 313
Jaycees, the (Junior Chamber of
Commerce) 81
Jennings, Ann L. 253, 267
Jevons, William Stanley 117, 131, 135, 320, 335
Ji, Minsun 493
Jobs, Steve 24n16

Jochimsen, Maren A. 256, 263, 267
Johannes, Gert-Jan 270, 288
John, Arielle 20n12, 33
Johnson, Brian 39, 51
Johnson, Erin M. 449, 472
Johnson, Ron 503, 504, 505, 507, 519
Johnson, William R. 573, 574, 579
Jordan, Michael 309
Joseph, Craig 122n7, 130, 134
Jost, Lawrence 59n11, 74
Joyce, Richard 120, 121, 126, 127, 135
Judaism 79
judiciary 85, 333
judgment 40, 55, 58, 59, 67, 85, 91, 118, 120–3,
129, 131, 167, 209, 214, 215, 217, 279,
295, 301, 309, 312, 342, 389, 412, 479,
481, 483, 485, 487, 488, 515, 530,
532–7, 545, 552, 583, 584
and willpower 60–2
Just, Richard E. 402, 422
just society 6, 139, 155, 382, 560
labor markets in 391–5
justice 1, 5, 7, 18, 18n8, 24, 24n17, 38, 39, 49, 65,
79, 83, 86, 90, 98, 99, 106, 118, 120,
123, 128, 139, 143, 145, 154, 155, 217,
220, 232, 233, 248, 250, 251, 255–7,
264, 265, 296, 301, 302, 308–10,
312, 355, 365, 366, 392, 425, 438,
488, 489, 508
arises from conflict 545
before 547–50
capabilitarian theories of justice 96
conceiving as realistically ecological 545–7
ethics and economies of ecological 544–57
manages traffic 545
pertains to process 546
and rights 560–1
social 96, 327, 374

K

Kabeer, Naila 257, 267
Kagan, Shelly 521, 522, 523, 524, 525, 526, 528,
530, 540
Kahlenberg, Richard D. 488, 494
Kahneman, Daniel 502n3, 518, 519
Kajackaite, Agne 187, 206
Kaldor, Nicholas 411, 422, 475, 480, 487, 494

- Kaldor-Hicks efficiency 66, 67, 67n27, 68,
 84–5, 90, 411, 474, 475, 478, 478n3,
 479, 479n4, 523, 524
 and impossibility of interpersonal
 comparisons 411–13
 normative deficiencies 480–5
 utilitarianism and 66–7
- Kamm, F.M. 56n3, 74, 507n6, 519, 521, 524,
 526, 528, 541
- Kanbur, Ravi 479, 480, 494
- Kandil, Feriel 438, 443
- Kant, Immanuel 2, 4, 74, 79, 84, 85, 91,
 176n10, 249, 267, 305, 349, 350, 357,
 364, 367, 377, 480, 508, 525, 529
 deontology and Kantian ethics 55–9
 Kantian approach to economics 54–73
 Kantian duties in models of economic
 choice 59–62
 applications of 62–5
 Kantian implications for welfare
 economics, policy, and
 markets 65–72
- Kaplow, Louis 118n3, 135, 416, 422, 479, 480,
 494, 529, 533, 541
- Karni, Edi 438, 444
- Karolyi, George A. 270, 288
- Katz, Lawrence 385, 396
- Kauffman, Stuart A. 148, 158
- Keeney, Ralph L. 432, 444, 464, 472
- Kelleher, J. Paul 507n6, 519
- Kelman, Steven 536, 541
- Kelo v. City of New London* 85
- Kemp, Paul 493
- Keren, Gideon 523, 541
- Keynesianism 231
- Keynes, John Maynard 233, 277, 286–7, 288,
 383–4, 396, 477, 487, 494
- Khalmetski, Kiryl 189, 206
- Khouri, Andrew 388n4, 396
- King, Gary 477n2, 494
- Kindleberger, Charles P. 362, 377
- Kinney, David 471
- Kirzner, Israel Meir 71n30, 74, 239,
 240n3, 246
- Kitcher, Philip S. 142n3, 147, 158, 566, 567,
 578, 579
- Kittay, Eva Feder 251, 264, 267
- Kiwanis, the 81
- Klamer, Arjo xviii, 5, 86, 91, 94, 270–87, 288
- Kletzer, Lori G. 573, 578
- Knapp, Georg Friedrich 360, 377
- Knight, Frank H. 85, 89, 94
- Knijn, Trudie 250, 267
- Knobloch, Ulrike xviii–xix, 5, 74n20,
 248–65, 267
- knowledge 17, 19, 20n12, 22, 28, 36, 58, 106, 125,
 141, 143, 145, 147, 150, 154, 169, 172,
 173, 199, 231, 233, 235, 235n2, 237, 238,
 240, 243, 244, 273, 276, 277, 281, 285,
 297, 298, 307, 312, 325, 350, 351, 369,
 375, 401, 436, 506, 513, 517, 561, 597
- Knudsen, Thorbjørn xviii
- Kochanek, Kenneth D. 86, 94
- Kocher, Martin G. 194, 206, 437, 444
- Kolers, Avery 374, 377
- Kolgomorov-Smirnov test 184
- Kolm, Serge-Christophe 438, 444
- Kolm-Pollak subfamily 420
- Komlosy, Andrea 252, 267
- Kornai, János 349, 352, 357
- Kornhauser, Lewis A. 521, 535, 541
- Koslowski, Peter 367, 369, 370, 377
- Krasemann, Keith W. 130, 137
- Kremer, Michael 448, 472
- Kremer, Monique 250
- Kronman, Anthony T. 67, 74
- Kropotkin, Petr A. 126n11, 135
- Krueger, Alan 502n3, 519
- Krueger, Anne 333, 335, 385, 389, 390, 391,
 396, 475
- Kuiper, Edith 253, 267
- Kuklys, Wiebke 110, 112
- Kumar, Nirmalya 14, 27, 31
- Kupers, Roland 155, 156
- ## L
- labor
- law 386, 388, 395
 - markets
 - ethics and, in, for 381–95
 - in general 567
 - in just society 391–5
 - and legitimate wages and involuntary
unemployment 382–5

- questions and answers about 567–76
 and rise of gig economy 385–91
 what do economists know about
 race in 566–76
 power 321, 381
 saving bias 322
 vs. work 252
 Labor Theory of Value (Marx) 321
 Lacke, Jay C. 367, 378
 laissez-faire 231, 329, 332, 334
 liberalism 332
 Lal, Deepak 166, 177
 Landes, David 309n14, 314
 Landes, William 535, 541
 Landsburg, Steven E. 41, 51
 Lane, Harlan 502, 519
 Lane, Robert E. 284, 288
 Lane, Ruth 140, 158
 Lang, Kevin 574, 579
 Lange, Oskar 325, 335, 338–9, 353, 357
 Langrill, Ryan 29, 31
 Langtry, Bruce 374, 378
 language acquisition 129
 large group cooperation 161–4, 173, 176n11
 large-scale society
 diversity of compliance behavior 143–5
 size of rule system 145–6
 Larmer, Robert 374, 378
 Lavoie, Don 281, 288
 law 6, 16, 29, 56–65, 70, 71, 80, 83, 85, 89, 121,
 122, 146, 193, 225, 238, 239, 321,
 328, 333, 341, 342, 343, 349, 382,
 386, 388, 390, 393, 395, 480n5,
 482, 549, 553, 555n11, 559, 561,
 563, 564, 568, 575
 deontological morality and economic
 analysis of 521–38
 economic approach to 64
 normative economic analysis of 522–4
 respect for 121
 when are laws needed? 563–4
 Law of Tendency for the Rate of Profit to
 Fall (Marx) 321
 Lawson, Robert A. 308n11, 313, 370, 378
 Leach, John 440, 444
 learning 14, 90, 128, 129, 130, 140, 150–3, 168,
 263, 277, 433, 439, 478
 Leemans, Inger 270, 288
 Leeth, John D. 567, 579
 legal regulation 146
 legitimate wages
 and labor markets and involuntary
 unemployment 382–5
 Leibbrandt, Andreas 189, 206
 Leiser, David 190, 207
 Lempert, Robert J. 441, 444
 Lerner, Abba 325
 Levine, Paul xv
 Levinson, Stephen C. 129n16, 133
 Levitt, Steven 571, 578
 Levy, David M. 18n8, 31, 85, 94
 Lewis, Paul 140, 146, 147, 158
 liars, types of 191–2
 Libertarian Paradox 469–70, 471
 liberty 49, 92, 99, 130n17, 145, 146,
 149, 232, 298, 319, 320, 322,
 480n5, 561
 Liedekerke, Luc Van 370, 377
 Lien, Tracey 388n4, 396
 life expectancy 86, 316n1, 441, 503
 Linarelli, John 371–2, 378
 Lindblom, Charles 330n11, 335
 linguistic communication 125
 Lions, the 81
 Lippert-Rasmussen, Kasper 561, 579
 Lipsey, Richard G. 337, 357
 litigation 333, 483n8
 Loewenstein, George 14n4, 30, 37, 50
 Lomasky, Loren 511, 512, 519
 Long, A.A. 48, 51
 Long, John D. 364, 378
 Lopes, Helena 132, 135
 loss frame 192
 lottery 84, 402, 423, 424, 428, 429, 431,
 434, 564
 Loveman, Gary W. 14, 27, 31
 Löwel, Siegrid 167n8, 177
 Lowry, James 366, 376
 Lowry, Rosemary 530, 532, 536, 541
 Lucas, Robert 383, 396
 luck egalitarianism 311
 Lutz, Mark 44, 51
 Luxembourg, Rosa 260, 268, 435, 438
 Lydenberg, Steven 368, 369, 378

lying

- aggregate 184
- aversion to 184
- behavior
 - in die-in-the-cup paradigm 184–5
 - in sender-receiver games 183–4
- cost of 188–9
- cross-country differences in 191
- disadvantageous 186
- evidence on 183–7
 - experimental methods 183–5
 - external validity of lab evidence 186–7
 - from field 185–6
- in gain and loss frames 192–3
- gender differences in 191
- in groups 194–5
- heterogeneity of behavior 191, 198
- history-dependent 194
- incentives and 187
- individual 184
- partial 185

Lynch, Tony 338, 357, 363, 380

M

- MacCallum, Gerald G., Jr. 100, 112, 511n8, 519
- Macfie, A.L. 14n5, 20, 31
- Machan, Tibor R. 370, 378
- Machina, Mark J. 428, 444
- MacIntosh, Alastair 365, 380
- MacIntyre, Alisdair 29, 31, 349, 357
- Mackenzie, Craig 374, 378
- MacKenzy, Donald 276, 282, 288
- Mackie, J.L. 120, 135
- Macleod, Henry Dunning 360, 378
- Macneil, Ian R. 14, 28, 31
- macro adaptations 148–50
- macro selection 148–9
 - evaluating upshot of 149–50
 - when pressures ease 150–1
- macro-level poverty analysis 105
- macro social evolution 148
- macroprudence 371
- Madoff, Bernie 563
- Madörin, Mascha 263, 264, 268
- Magala, Slawek 270, 281, 288
- Maitland, Ian 364, 378
- Malcai, Ofer 534, 535, 541

male breadwinner and female housewife
model 260

male-centered ethics *see* androcentric ethics

male ethics of justice 250

malice 198

Malthus, Thomas Robert 123n9, 144n9, 158

Malthusianism 144

managerial power 392, 393

managers 80, 282, 339, 343, 344, 352–4, 367,
368, 392, 394

Mandeville, Bernard 34, 35–6, 38, 39, 40, 52,
296–7, 296nn2,3, 299, 306, 314

Maniquet, François xvii, 423, 443

Mankiw, Gregory 360, 378

Manne, Henry G. 370, 378

Manove, Michael 574, 579

Mao Zedung 278

marginal analysis 64, 65

marginal productivity 332

marginal utility 230, 411, 414–16, 425, 427, 440

Marglin, Stephen A. 483, 494

Margolis, Howard 2, 8, 63n18, 75, 130, 135

Marino, Patricia 92, 94

market

- adversarialism 348–54
- condition 201, 338
- cost-effectiveness of relying on 503–7
- economies, evolution of 160
- equilibrium 35, 64
- exchange 14, 64, 237, 286, 300, 305
- failure 13, 70, 70n28, 71, 88, 559
- nature and regulation of 70–2
- power 345, 348, 354
 - exercising 345
- socialism 236
- sphere 104, 273
- transactions 14, 216, 217, 300, 303, 306,
370, 523
- values 210, 215, 216
- volatility 367, 487

marketization 329

markets

- Adam Smith on ethics and 20–8
- Adam Smith's theory of 15–18
- as adversarial institution 350
- can markets be humane? 310–13
- and commerce 293–395

- dehumanizing 304–6
 humane 295–313
 and incentives and morality 200–2
 Kantian implications for 65–72
 modeling markets and human behavior
 80–1
 noxious 307
 vice of 306–8
- marriage 2, 63, 64, 394, 513, 562
 laws 250, 565
- Marshall, Alfred 233, 234, 235
- Marshall, Thomas Humphrey 381, 396
- Martin, Dominic 350, 357
- Martin, Felix 360, 378
- Marx, Karl 5, 25n18, 31, 123n9, 268, 277, 278,
 283, 288, 295, 295n1, 304, 305, 306,
 306n8, 314, 320–2, 332, 336, 339, 363,
 378, 381, 396, 491
 on capitalism and democracy 320–2
- Marxism/Marxian economics 276, 283, 318,
 323, 324, 330, 332, 491
- Maseland, Robbert 270, 274n2, 276n3,
 278n4, 287
- mass unemployment 233, 237
- material dimensions 104
- materialism 319
- mathematics, uses of 599–603
- Matthews, Peter Hans 128n15, 133
- Maurin, Eric 431, 443
- Max-C* 165, 168–73, 175, 176
- Max-U* 4, 160, 164–5, 170–6
 bookshelf metaphor 165–70
 neoclassical emphasis on, and classical
 emphasis on conformity 164–5
- Mayhew, Anne 253, 268
- Mazar, Nina 187, 188, 206
- McCabe, Douglas M. 367, 376
- McCabe, Kevin A. 81, 94
- McCall, John J. 364, 378
- McCarthy, David 596, 597, 598, 599,
 604, 605
- McCarty, Richard 363, 378
- McCloskey, Deirdre N. xvii, 2, 3, 4, 7, 8, 14, 18,
 29, 31, 38, 49, 52, 88, 91–2, 93, 94,
 132, 135, 160, 177, 240, 246, 270, 273,
 274, 277, 278, 280, 281, 283, 285, 289,
 308n11, 309n14, 314, 482n6
- McCrink, Kileen 128, 135
- McCurdy, Howard E. 70n28
- McGowan, David 482, 494
- McIntyre, Alison 525, 541
- McMaster, Robert xvii, 222n5, 227
- McNally, Peter 140–1, 156
- McNaughton, David 528, 532, 541
- McPherson, Michael S. xviii, xix, xxi, 3, 6, 8,
 36, 51, 138, 157, 381–95, 396
- mechanism design 182
- Medema, Steven G. 54n1, 65, 73, 75
- medical care 307, 499, 504
- medical ethics 97, 475
- Medina, Barak xix–xx, 6, 55n2, 61n15, 65n21,
 520–38, 543
- Mediterranean economy 331
- Meehan, Jennifer 277, 288
- Meikle, Scott 36n2, 44, 52
- Mendelson, Tal 538
- Menger, Carl 117, 136, 320, 336, 359, 378
- mental-state accounts 401
- Menzel, Paul 500n1, 519
- mercantilism 333
- Mercuro, Nicholas 65, 75
- Merkel, Wolfgang 331n13, 336
- Mesoudi, Alex 148, 158
- metachoice 62
- metallism 360
- methodological individualism 130, 253
- Meyer, Marco 373, 378
- microcredit 372
- microfinance 265, 371–3
- Microsoft Corporation 341
- Midgley, Mary 126, 136
- Mies, Maria 260, 268
- migration 124, 125, 150, 251, 263, 265,
 333, 440
- Miguel, Edward 448, 472
- Mikhail, John 526, 541
- Mill, James 232, 305
- Mill, John Stuart 79–80, 82, 86, 94, 147, 158,
 240, 393, 394, 396
- Miller, Dale E. 418, 421
- Miller, David 139n1, 158
- Miller, John H. 148, 158
- Millgram, Elijah 557, 558
- Millner, Antony 436, 444

- mind 19, 43, 48, 82, 83, 142, 163, 166, 171, 174,
175, 501, 511, 546
and rules 140–1
mind game 187
Minimal Libertarianism, Principle of 469
minimum wage laws 383, 388, 389
Minkler, Lanse P. 59, 75, 132, 136
Minsky, Hyman 362, 378
Mirowski, Philip 277, 289
Mirrlees, J.A. 424, 444
mirror neurons 127n14
miserliness 41
Mises, Ludwig von 230, 231, 235, 236, 240,
241, 242, 243, 244, 246, 325,
328, 336
Mishkin, Frederic 362, 378
missing markets 70
Mitchell, Melanie 147, 147n10, 158
Mitchell, Robert C. 534, 541
modeling distribution sensitivities 452–4
moderate deontology 522, 529, 535
and absolutist deontology 526–7
and the law 530–1
modified breadwinner and carer model 260
Mohanty, Chandra Talpade 251, 268
monetarization 261
monetary equivalent 400, 403, 404, 412, 414,
415n20, 420
money 1, 5, 23, 42n5, 44, 59, 60–3, 71, 80, 83,
86, 101, 105, 175, 183, 190, 197, 217,
218, 220, 224, 230, 241, 252, 273, 274,
276, 277, 284, 286, 287, 317, 318, 328,
329, 333n14, 334, 342, 384, 388, 395,
403, 411, 412, 415, 425, 427, 448, 457,
479, 482, 500, 524, 527, 587
commodity theory of 359, 360
credit theory of 360
and finance, ethics of 358–76
as root of all evil? 363–7
state theory of 360
what is it? 359–61
Mongin, Philippe 404, 420, 422, 426, 427, 428,
438, 442, 443, 444, 593, 597, 605
monopoly 85, 235, 323, 333, 344
monotonicity axiom 405n7
Montes, Leonidas 14n5, 18n8, 31
Moore, Jennifer 369, 370, 378
Moore, Michael S. 61, 75, 526, 530, 541
moral balancing 189–90
moral behavior 29, 46, 147, 183, 306
explanations for 188–91
cost of lying 188–9
maintenance of self-image 188
moral balancing 189–90
public personas, private personas, and
morality 190–1
self-serving justifications 190
moral claims 35, 45, 121, 129, 255
moral codes 123, 125
moral development 21, 26–9, 128, 131, 250
moral disengagement 200
moral harm 251, 264
moral ideals and political compromise 556
moral judgments 55, 118, 120–2, 129, 131,
530, 532–5
moral limits of markets 212, 213, 220, 224–5
moral motivation and evolution in
economics 117–32
moral preferences 173, 223, 533–5
moral relativism 122, 122n6
moral restraint, duty-based 173–7, 176n10
moral risk 26
moral sentiments 4, 14, 65, 77, 80, 81, 90, 91,
117, 118, 125
moral space, eligible and types of rules 144
moral standpoint of care 249
moral status of profit 337–56
moral suasion 183, 191, 352
and morality 195
moral systems 21, 127, 129, 130, 132
moral values 121, 123, 165, 219, 298
and capabilities and human dignity
221–4
and four different types of individual
advantage 222
positions regarding pervasiveness of, in
economic life 224
morality 4–6, 13, 15, 29, 49, 56, 58, 78, 181–3,
187, 192, 193, 201, 202, 239, 242, 254,
304, 312, 337–9, 349, 350, 353, 354,
363, 366, 549
and anonymity 197
and beliefs: psychological game theory
198–200
as complex adaptive system 138–55
Darwin and evolution of 123–5

- deontological, and economic analysis of
law 521–38
how it evolved 127–31
and incentives and markets 200–2
and moral suasion 195
more than skin deep 125–7
public personas, private personas,
and 190–1
and social identity 196
sources of, in economics 117–77
what is it? 119–23
morally irrelevant costs and benefits 527–8
Morduch, Jonathan 372, 373, 376, 377
Moreau, Sophia R. 536, 541, 561, 562, 565,
568, 579
Morgan, John 308n11, 313
Morgenstern, Oskar 587, 588, 606
Morrow, Glenn R. 18n8, 31
mortgage loans 340
Mosher, Dave 86, 94
motivation 5, 13, 36, 37, 42n5, 44, 46, 47, 51, 63,
64, 182, 202, 239, 297, 300, 337, 364
moral, and evolution in economics 117–32
extrinsic 200
intrinsic 200–1
motor behavior 142
muddling through 93
Muellbauer, John 585, 605
Mukerjee, Swati 567, 579
Muldoon, Ryan 139, 159
Mulgan, Tim 525, 541
Mullainathan, Sendhil 196, 206, 570, 571, 578
Muller, Jerry Z. 14, 17, 26, 28, 31
multidimensional analysis 99
multidisciplinary work 209
multilateral setting 201
multiple self 130
multiple utilities 62
Munger, Michael 46, 52
murder 55, 58, 121, 122, 170, 171, 528
Murray, Charles 573n5, 579
Murray, Christopher 503n4, 519
mutual sympathy of sentiments 301
Myrdal, Gunnar 229, 246, 566n2, 579
- N**
Nagel, Thomas 466, 472, 524, 525, 527, 541
Natali, Carlo 46, 52
National Longitudinal Study of Youth
(NLSY) 573
nationalism 298, 299, 316
natural liberty (Smith) 145, 146, 232
Nazism 328
Neal, Derek A. 573, 574, 579
Neanderthals 548–49
necessary vs. unnecessary harms 487
negative externalities 201, 202, 217, 308
negative reciprocity 164
Nelson, Julie A. 2, 8, 81, 94, 253, 254, 256, 257,
262, 268
neoclassical economics 13, 37, 82, 160, 181,
181n1, 182, 191, 192, 214, 231, 253,
253n6, 259, 263, 325
neoclassical general equilibrium analysis 117
neoclassical welfare theory 82–3, 84
neoliberalism 253, 256, 257, 278, 331n13
Netherlands, the 507, 514
neural mirroring 127n14
neuronal network 140
neuroscience 126, 167n8
New Deal 329
Neysmith, Sheila M. 262, 268
NHE (New Home Economics) 253n6, 259
Nichols, Shaun 122n7, 136, 143, 144, 145, 157
NIE (New Institutional Economics) 253n6
Noddings, Nel 251, 268
nomein 272
non-aggression axiom 239
non-compensable vs. compensable harm
481–3
non-consumption attributes 403, 411n18,
415n20
non-tuism 175, 241, 243
Norcross, Alastair 525, 542
Nord, Erik 508, 519
Nordenfelt, Lennart 500n2, 519
normative economics 78, 81, 92, 522, 527, 535
analysis 522–4, 533, 535
of social risk 423–41
normative moral relativism 122
normative negligence 474
North America 233
Northern Europe 279, 280
Norton, Bryan 545n1, 558
noxious markets (Satz) 307
Novak, Michael 14, 31

Nozick, Robert 106, 112, 387, 396, 401, 422,
489n11, 494
NPOs (non-profit organizations) 257, 261, 262
Nussbaum, Martha 4, 42, 52, 86, 96, 98n1,
101, 106, 112, 256, 268, 308n12, 314,
483, 488, 489, 494

O

O'Donoghue, Ted 62n17, 75
O'Donovan, Carolyn 387n3, 390n6, 396
O'Flaherty, Brendan xix, 7, 559–78, 579
O'Neill, Martin 359n1
O'Neill, Onora 58, 75
obedience 123, 124, 171
Oberholzer-Gee, Felix 81, 94
objective-good accounts 401
objective list theories 104
obligation 24, 58, 61n14, 65, 71, 72, 87, 186,
302, 317, 339, 341, 342, 351–3, 351n3,
355, 356, 362, 366, 368, 369, 373, 375,
387, 517, 524
Obrinsky, Mark 339, 357
OECD 105
Oehmke, Martin 371, 372, 376
oikeiosis 46
oikonomia 251–2
oikos 272
Okasha, Samir 148, 158
Okin, Susan Moller 251, 258, 268
Olivier, Jacques 284, 289
Olson, Mancur 334, 336
Omtzigt, David xv
Oppenheim, Felix 106, 113
Oppenheimer, Stephen 127, 136
opportunism 57, 64, 89, 152, 153, 161–4, 168,
169, 172–5, 365, 370
oppression 130, 130n17, 233, 253, 322, 564
ordoliberalism 256, 257
organizational objective, profit as 337, 343
organizational theory 139
Ortiz-Ospina, Esteban 175n9, 177
Osteen, Mark 270, 290
Ostrom, Elinor 146, 158
other-regarding preferences 4, 60, 63, 90, 118,
120, 130, 131, 164, 181n1, 182, 184, 191
Otsuka, Michael 461n7, 466, 467, 472, 588,
589, 597, 598, 605

objection to prioritarianism 467–8
Otteson, James R. xix, 5, 14n5, 17, 20, 21, 31,
177, 295–313, 314
“ought” type statements 120, 214, 215, 219,
220, 412

P

Pack, Spencer J. 26, 31
Paganelli, Maria Pia 14, 14n5, 21, 28, 31, 32, 177
Page, Scott E. 139, 147n10, 148, 155, 158
Pager, Devah 570, 578
paid and unpaid work 259–62
economic theory of 261–2
pain 18, 48, 59, 318, 484, 529, 599, 600
and pleasure 78, 82, 91, 418, 523
Palmer, Craig T. 125, 136
parent–child relationships 111
parental care 131
Pareto, Vilfredo 83–4, 94
Pareto
axioms 400, 405n7, 411n19, 432, 440
criterion 67, 110, 523
utilitarianism and 66–7
efficiency 83–4, 423
indifference 400, 403, 410, 410n14, 412
vs. inequality aversion 429
optimality 82–4, 216, 348, 423, 523
principles 400, 410, 411, 417, 419, 420, 425,
426, 427, 429–31, 433, 439, 469, 523,
528, 529, 594
strong 400, 403, 405n7, 410, 410n14,
411n19, 412
superiority 412, 413
test 67, 84, 479
Parfit, Derek 90–1, 95, 104, 113, 401, 401n4,
418, 419, 422, 434, 444, 467, 468,
472, 501, 509, 519, 532, 541, 589, 590,
591, 592, 593, 601, 603, 606
Parkin, Michael 360, 378
patent protection 85
Paton, H.J. 75
patriarchal dividend 250, 257, 258
Pattanaik, Prasanta K. 110, 113, 438, 444
Paul, Ellen Frankel 139n1, 158
payments system 361
Payne, John W. 534, 542
Paynter, Brook S. 575, 579

- Payutto, Prayudh A. 279, 289
 Pearse, Rebecca 250, 253, 268
 Peart, Sandra J. 18n8, 31, 85, 94
 Peer, Eyal 190, 206
 peer group interaction 131
 Peil, Jan 3, 8
 perceived cheating aversion 200
 perception-action cycle 161
 perfect duty 60, 63
 Peripatetic school of philosophy 47
 permissive rule systems 144
 Person Separability 417–19
 personality variance 142
 Peterson, Martin 530, 532, 536, 541
 Petrie, Ragan 196, 204
 Pettifor, Ann 361, 378
 Pezzullo, Phaedra 545n1
 Pfattheicher, Stefan 192, 206
 Phelps, Edmund 36, 52
phronesis 91
 Pigou, A.C. 68, 75, 230, 234, 235, 236, 246
 Pigouvian taxes 68, 70
 Pigou-Dalton axiom/condition 419, 427,
 589–90, 591, 592, 593, 595
 Piketty, Thomas 308n12, 309, 314
 Pilbeam, Keith 362, 378
 Pinker, Steven 129n16, 136
 Pitts, Jennifer 14n5, 20, 32
 Pivato, Marcus 420, 422, 426, 428, 438, 444
 plagiarism 88
 Plato 154, 365, 378, 550
 Euthydemus 39
 Plattner, Marc 316, 331, 335
 pleasure 79–80, 82, 90, 117, 253, 298, 301, 561
 and pain 78, 82, 91, 401, 418, 523
 Ploner, Matteo 190, 206
 Plonz, Sabine 254n7, 256, 256n8, 268
 plundering 305
 pluralism in economics 77–95, 99
 arguments against 89–2
 relativism 91–2
 ultimate ends 89–90
 unifying across frameworks 90–1
 ethical 77–93
 horizontal and vertical 78–80
 Plutarch 48
 Polania-Reyes, Sandra 200, 203
 Polanyi, Karl 5, 252, 268, 271, 284, 289,
 328–30, 330n9, 331, 332, 334, 336
 on capitalism and democracy 328–30
 policymaking 2, 6, 65, 105, 414, 522, 531
 political compromise and moral ideals 556
 political donations 342
 political economy 3, 5, 14, 15, 29, 37, 191, 230,
 232, 233, 235, 239, 252, 256n8, 259,
 316n1, 320, 330, 416
 classical tradition of 295–313
 criticisms of Hume and Smith 303–10
 humane 296–303
 political fraud 198
 political science 211, 211n1
 political theory of care 250
 pollution 68, 70, 387
 Ponzi schemes 367
 Ponthière, G. 441
 Poor Law (UK) 29
 Portmore, Douglas W. 532, 541
 positive-normative distinction 213–15, 218,
 219, 221, 226, 227
 positive reciprocity 164
 positivistic empiricism 231
 Posner, Eric A. 402, 414, 421, 479, 479n4,
 480, 488, 493, 521, 532, 533, 534, 535,
 538, 541
 Posner, Richard A. 58, 69, 75, 524, 541, 542
 postcolonial studies 258
 postwar literature on capitalism and
 democracy 330–4
 crises of democratic capitalism 332
 rent-seeking 333–4
 varieties of capitalism 330–1
 poverty 98, 99, 105, 110, 232, 233, 235, 282,
 309, 310, 316n1, 324, 372
 Powell, Russell 129, 133
 Power, Marilyn 262, 268
 power and compulsion 329
 pragmatism 91
 Prainsack, Barbara 509, 520
 Pratt, Andy 270, 283, 289
 preference rankings 62, 131n18
 preference satisfaction 2, 54, 55, 83–6, 89, 215,
 412, 488, 501, 523, 532
 prestige-based imitation 128
 prestige bias 151

- prevalence-of-rule-violations (PRV) index 198
- Price, Michael E. 128n15, 136
- price system 237, 240, 244, 325, 349, 474
- pricing decisions 354
- primatology 118
- prioritarian SWFs 400, 405n9, 407, 411, 418, 419, 420
- prioritarianism 6, 7, 418–20, 434, 468, 589–599, 604
- bad argument for 592–3
- definition of 589–91
- good argument for 593–6
- Otsuka and Voorhoeve's objection to 468
- prisons 385
- private provisioning 262
- Priyadarshree, Anurag 373, 378
- production 2, 40, 80, 108, 130, 143n8, 172, 202, 230n1, 231, 233, 236, 237, 239, 240, 252n5, 253, 259, 261, 263, 264, 276, 281, 283, 284, 303, 305, 321, 323, 325, 330n9, 333, 339, 348, 352–4, 383, 388, 389, 491, 553
- and exchange, microeconomic models of 64
- means of 352
- productivity 46, 163, 164, 241, 242, 243, 252, 264, 277, 298, 332, 383, 395, 486, 501, 506, 514, 551, 555, 563, 572, 573
- professional identity 196
- profit 5, 24, 37, 42n5, 46, 54, 63, 80, 84, 85, 141, 252, 257, 279, 302, 307, 321, 363, 364, 366, 373, 510n7, 559, 577
- maximization 337, 348
- moral status of 337–56
- as organizational objective 337, 343
- what is it? 339–43
- profit-seeking 279, 302
- why it is collectively self-defeating 343–8
- prohibitions-based social policies 224–5
- prohibitory rule systems 144
- property rights 24, 71, 72, 85, 106, 234, 236, 240, 253, 299, 317, 318, 320, 322, 324, 330n9, 332, 333, 553, 555
- prosperity 5, 7, 25, 231–3, 240–4, 295, 297, 298, 299, 302, 303, 308–12, 308n11, 332, 362, 499
- protecting strata of society (Schumpeter) 323
- protectionist international trade policies 333
- Proudhon, Pierre-Joseph 318
- Protestantism 83, 197, 279, 280, 364, 365
- provisioning sovereignty 262
- provisioning work 262, 264
- individual 262
- social 262
- Pruckner, Gerald J. 186, 206
- prudence 18, 21, 24, 79, 311, 371, 490
- psychological game theory (PGT) 198–200
- psychology 36, 38, 39, 47, 58, 126, 142, 152, 188, 211n1, 301
- public choice theory 334, 533
- public good 88, 127, 200, 277, 364, 416, 479, 485, 531, 534, 576
- public personas, private personas, and morality 190–1
- public policy 83, 84, 86, 91, 92, 155, 235, 236
- public safety 71, 79, 86
- public sphere 77, 190, 191, 531n2
- public transport 109, 186
- “public benefit” (Mandeville) 296, 297
- Pulford, Briony xv
- Pullman, George 394
- punishment 65, 121, 128, 173, 183, 312, 490
- and delegation 193
- pro-social 193
- third party 193
- purpose-independent rules 149
- Putnam, Hilary 219, 220, 228
- Putterman, Louis 128n15, 132
- ## Q
- Qizilbash, Mozaffar 104, 113
- quality of life 99, 105, 600
- queer ethics 251
- Quinn, Robert 270, 281, 287
- ## R
- Rabin, Matthew 62n17, 75
- Rabinowicz, Wlodek 597, 598, 599, 604, 606
- race 252, 537
- civil rights, and employment 558–78
- in labor market, what do economists know about 566–76
- Radin, Margaret Jane 307n9, 314
- Ragatz, Julie 369, 379
- Ramsey, Frank 587, 606
- Ramstad, Yngve 491, 494

- randomization 437, 447, 455, 456
- Rao, Vijayendra 270, 281, 282, 289
- Raphael, D.D. 18, 18n8, 24n17, 32
- Rasmussen, Dennis C. 14, 14n5, 17, 18, 20, 23n15, 29n19, 32
- rational choice model 164, 172
- rationality 13, 14, 37–40, 42, 44, 45, 47, 48, 151, 182, 191, 250, 252n5, 254, 258, 318n3, 323, 323n5, 368, 401, 402, 423, 425–9, 432, 435, 436, 439, 441, 451, 509
and ethics and economic behavior,
evolution of 160–77
flaws in neoclassical conception 14
- Ratcliffe, Susan 89, 94
- Ravallion, Martin 478, 494
- Rawling, Piers 528, 532, 541
- Rawls, John 66, 75, 139, 158, 256, 268, 396, 438, 444, 446n1, 473, 489n11, 494, 508, 520, 521, 524, 542, 560–1, 579, 591, 606
- Razavi, Shahra 260, 261, 263, 268
- Read, Leonard E. 16n6, 32
- reciprocity 37, 81, 123, 125, 164, 182, 198–200, 303, 350, 356
golden rule of 249
indirect 36
sphere of 271
strong 128
- redistribution 83, 191, 235, 271, 282, 309, 331–4, 412, 413, 425, 439
- Reflexionsstopp* (abandonment of reflection) 254
- reflexivity 251, 362
- Reformation 553
- Regner, Tobias 190, 206
- regression analysis 281
- regulation 1, 70–2, 88, 92, 107, 108, 146, 195, 196, 235, 256, 273, 277, 282, 307, 329, 331–4, 338, 343, 351, 369, 375, 383, 386–8, 390, 391, 416, 440, 476, 487, 491, 505, 507
- Rehavi, M. Marit 449, 472
- Reibetanz, Sophia 461n7, 463, 473
- Reich, Robert 334, 336
- Reid, Margaret G. 252, 268
- reinforcement learning 129
- Reinhart, Carmen M. 362, 379
- Reinhardt, Uwe E. 86, 95
- Reiss, Julian xix, 5, 316–34
- relativism 90–2
moral 122, 122n6
- religion 83, 89, 121, 124, 173, 271, 273, 279, 286, 305, 320, 524, 537, 545, 562
economic research development as
akin to 89
- Renda, Andrea 399, 422
- Rendleman, Doug 482, 495
- rent 225, 319, 348
- rent-seeking 333–4
- renumeration 278, 285
- reparable vs. irreparable harms 487
- resource allocation 80, 85, 176, 234, 332
- restrictive intervention 447, 455–6, 460, 462
- Rethinking Economics 3
- Reuben, Ernesto 193, 206
- reverse engineering of evolved complex system 150
- Reznek, Lawrie 500n2, 520
- Rheinberger, Christophe M. 433, 444
- Rhonheimer, Martin 36, 52
- Ricardo, David 232, 321, 336
- Rich, Arthur 251n3, 255, 269
- Richardson, Benjamin J. 373, 379
- Richerson, Peter J. 124, 127, 128n15, 132, 142n4, 143n7, 148, 150, 151, 151n13, 158, 163n5, 177
- rights
care 249
civil 7, 190, 317, 327
see also civil rights laws
human 49, 72, 78, 79, 85, 222–4, 249, 373, 524, 560
property 24, 71, 72, 85, 106, 234, 236, 240, 253, 299, 317, 318, 320, 322, 324, 330n9, 332, 333, 553, 555
of refusal 67
“trump” welfare 68
wrongs and externalities 67–71
- risk 6, 107, 108, 163, 169, 193, 202, 209, 214, 223, 274, 320, 329, 340, 346–8, 358, 361, 365, 367, 368, 373, 387, 394, 499, 510, 512, 517, 530, 551, 586–8, 595, 597
moral 26
normative economics of social risk 423–41
systemic 371–2
-taking 85, 92, 280, 282, 283, 285, 366, 375

- risk (*cont.*)
 unwelcome reductions 448–9
 and sensitivity to intrastate
 distribution 457–8
 and welfare and policy 397–493
 risky decisions, ethics of making for others
 446–71
 Risse, Mathias 564, 579
 Rizzolatti, Giacomo 127n14, 136
 Robbins, Lionel 35n1, 52, 54, 54n1, 75, 139, 158,
 214, 214n2, 227, 228, 230, 231, 235,
 236, 244, 246, 276, 289, 295, 314,
 412, 422
 Roberts, Jessica L. 536, 542
 Roberts, Peter W. 14, 32
 Roberts, Simon 122n7, 136
 Robertson, Dennis H. 32, 239, 246
 Robeyns, Ingrid xx, 4, 96–111, 113, 256,
 256n9, 269
 Robinson, Fiona 250, 269
 Robinson, Jean 231
 Robinson, Paul H. 128, 136, 158, 531, 542
 Rodgers, William, III 574, 579
 Roemer, John E. 322, 336, 349, 357, 438,
 440, 445
 Rogoff, Kenneth S. 362, 379
 Roman Catholicism 279, 280
 Roodman, David 373, 379
 Roosevelt, Franklin D. 329
 Rose, David C. xx, 4, 142n5, 159, 160–77, 178
 Rosenberg, John S. 85, 95
 Rosenberg, Nathan 23, 24, 32
 Rosenberg, Richard 373, 379
 Roser, Max 175n9, 177
 Ross, Don 36, 37, 52
 Ross, W.D. 75, 528, 542
 Rotary International 81
 Roth, Alvin E. 220, 224, 228
 Rothbard, Murray Newton 234, 239, 246, 361,
 378, 379
 Rotten Kid theorem (Becker) 64
 Rothschild, Emma 14, 14nn3,5, 17, 32
 Rousseau, Jean-Jacques 5, 296, 304, 305, 314
 rule-consequentialism 532
 rules 7, 21, 29, 39, 61, 78–80, 87, 91, 92, 101,
 103n2, 107, 108, 127–9, 140–55, 165,
 175, 190, 231, 238–40, 242–4, 250,
 252, 273, 302, 311, 312, 333, 343, 349,
 351, 368, 383, 387, 390–4, 439, 485,
 488, 516, 522, 523, 529, 536, 547
 and action 141–2
 Bodo 142–3
 moral rules 38n4, 118, 121–3, 129, 139, 142,
 145, 148, 149, 337
 size of system 145–6
 social rules 139, 142, 142n3, 145, 148, 152,
 153, 155, 240
 types of, and eligible moral space 144
 and mind 140–1
 violation 153
 Runciman, Walter Gary 130, 136
 Ruser, John 567, 579
 Russell, Daniel 36, 39, 46, 52
 Rustichini, Aldo 201, 205
 Rutland, Peter 352, 357
 Ryan, John A. 366, 379
- ## S
- Saari, Donald 153n15, 159
 Safri, Maliha 260, 269
 Sampson, Geoffrey 129n16, 136
 Samuelson, Paul 276, 289
 Sanchirico, Chris W. 426, 441
 Sandberg, Joakim xx, 5–6, 358–76, 377, 379
 Sandel, Michael J. 29, 32, 37, 42, 52, 220,
 224, 228, 270, 289, 295n1, 306, 307,
 307n9, 314
 Sandis, Constantine 490, 495
 Sandler, Ronald 545n1, 558
 Sap, Jolande 253, 267
 Satz, Debra xviii, xix, xx–xxi, 3, 6, 8, 111, 113,
 295n1, 307, 307n9, 314, 381–95
 Sausgruber, Rupert 186, 206
 Sautet, Frederic 71n30, 75
 Savage, Leonard J. 435, 445
 “Savage man” (Rousseau) 304
 saving few at greater risk to many 460–3
 savings 60, 285, 375, 516
 Say, Jean-Baptiste 232, 240
 Scandinavia 331
 Scanlon, Thomas 501, 507n6, 520
 Scheffler, Samuel 66n23, 75, 525, 530, 542
 Scheidel, Walter 310, 315
 Schelling, Thomas 461n7, 473

- Schindler, Simon 192, 206
 Schkade, David A. 534, 542
 Schleifer, James 318n4, 336
 Schlichter, Detlev 361, 379
 Schliesser, Eric 18, 18n8, 32, 36, 37, 52
 Schmeidler, David 436, 443
 Schmidt, Klaus M. 2, 8, 182, 191, 197, 205
 Schmidt, David xxi, 7, 240, 544–57, 558
 Schnabl, Christa 251, 269
 Schor, Juliet B. 307n9, 315
 Schroeder, Mark 532, 542
 Schudy, Simon 194, 206
 Schulz, Jonathan F. 198, 205
 Schumpeter, Joseph 5, 132, 136, 237, 246,
 318n3, 321, 322–4, 323nn5,6,
 325–7, 327n8, 328, 329, 332, 336,
 360, 379
 on capitalism and democracy 322–4
 Schwartz, Shalom H. 122n7, 136
 Schweitzer, Maurice E. 192, 206
 science and ethics 214, 220
 scientific inquiry 139, 164, 174, 239
 Scotland 280
 Scott, Allan J. 283, 289
 Scottish Enlightenment 550
 screening 6, 402, 448–9, 457–8, 470, 471,
 576, 577
 and no screening 457–8
 Scruton, Roger 275, 289
 Sean 448, 473
 Searle, John 360, 379
 Second World War 236, 330, 332
 Segal, Uzi 428, 442
 self-image 185, 188, 195
 self-interest 1, 2, 4, 6, 14, 17, 18, 20, 22n14, 28,
 36, 37n3, 39, 40, 45, 58–60, 64, 65,
 77, 80, 81, 89, 117, 118, 122, 126, 131,
 132, 142, 165, 172, 175, 186, 296, 297,
 299, 300, 337, 343, 364, 401, 402, 501
 self-love 14, 17, 21, 22, 300
 self-organization 155
 and return of invisible hand 153–4
 self-regarding preferences 181, 191, 192,
 215, 479
 self-regulating market system 328, 330
 self-respect 99, 479n4, 483, 484
 self-serving justifications 188, 190
 selfishness 20n13, 23, 29, 36, 42, 45, 57, 123,
 126, 126n12, 130, 174, 175, 197, 296,
 297, 300, 305, 363, 366
 axiom 42
 of genes 126
 Sen, Amartya 1, 4, 8, 13nn1,2, 32, 36, 45, 52, 59,
 63, 75, 81, 82, 86, 90, 95, 96, 97, 98n1,
 99, 101, 102, 104, 108, 109–10, 113,
 117n1, 119n5, 123nn8,9, 132, 136, 159,
 178, 221–2, 222n5, 223, 228, 239, 240,
 246, 256, 269, 270, 274, 280, 289,
 308n12, 315, 427, 445, 446n1, 473,
 479, 479n4, 480, 483, 489, 489n11,
 495, 523, 529, 532, 535n5, 542, 591,
 592, 597, 606
 Libertarian Paradox 469–70, 471
 sender-receiver games 187, 189, 198
 lying behavior in 183–4
 Seneca 43
 Sengupta, Sunil 97, 113
 sense-making 5, 270–87
 separability 426, 439, 590, 595
 Axiom of Weak Separability of
 Persons 465
 conundrum 434–5
 vs. ex post egalitarianism 434
 implications of 435
 Person-Separability 417–19
 Sethi, Rajiv xix, 578
 Sevenhuijsen, Selma 250, 251, 269
 sexual harassment 87
 sexual stereotyping 87
 Shahar, Dan xxi
 Shalvi, Shaul 190, 192, 206, 207
 shame aversion 198
 shareholder(s)
 as residual risk-bearers 340
 value 339, 341–3
 Sharon, Assaf 529, 540
 Sharpe, Matthew 47–8, 52
 Shavell, Steven 118n3, 135, 479, 480, 494, 529,
 533, 541
 shelter 14, 48, 104, 108, 242, 272
 Sherman, Nancy 59n11, 75
 Shiffrin, Seana V. 480, 483, 495
 Shiller, Robert 368, 379
 Shin, Patrick 560, 579

- Shrader-Frechette, Kristin 545n1, 558
 Siddhartha Gautama 79
 Siegfried, John 474, 475, 495
 Simler, Kevin 40, 52
 Simmel, George 284, 289
 Simon, Herbert 211–12, 211n1, 228
 Singapore 514
 Singer, Peter 531, 542, 545n1, 558
 Singer-Vine, Jeremy 387n3, 390n6, 396
 single cause 448, 456, 457
 and multiple cause 456
 Sinn, Hans-Werner 366, 379
 six worlds of value regimes (Boltanski
 et al.) 272
 Skidelsky, Edward 295, 299, 307n9, 315,
 338, 357
 Skidelsky, Robert 295, 299, 307n9, 315,
 338, 357
 skills 16, 25, 35, 286, 298, 308, 321, 331, 392,
 570, 573, 574n6
 Skeyms, Brian 151, 159
 slavery 225, 301, 559
 Sliwka, Dirk 189, 206
 slippery slope of dishonesty 197
 Smaga, Pawel 371, 372, 379
 Smart, J.J.C. 123n8, 136, 532, 542
 SMEs (small and middle enterprises) 257
 Smith, Adam 4, 5, 38n4, 39, 79, 81, 90, 91,
 92, 93, 95, 118, 145, 161, 164–7, 232,
 240, 245, 246, 249, 280, 296, 297,
 299–303, 304, 305, 306, 308–13,
 311n18, 315, 318–19n3, 336, 343, 359,
 364, 381
 butcher-brewer-baker passage 300
 Das Adam Smith Problem 14
 on ethics 20–9
 humane political economy 299–303
 and limits of benevolence 21–3
 and possibility of moral development
 26–8
 and potential for corruption 23–6
 and study of ethics in commercial
 society 13–33
 theory of ethics 18–20
 theory of markets 15–18
 works
 LI (*Lectures in Jurisprudence*) 24, 27, 32
 TMS (*Theory of Moral Sentiments*) 14–15,
 18, 19–22, 19n10, 20n11, 23, 24, 28,
 32, 80, 117, 123, 123n9, 159, 232, 269,
 289, 301, 303n7, 311, 312
 WN (*An Inquiry into the Nature and
 Causes of Wealth of Nations*) 14,
 15–18, 17n7, 20–1, 22, 23, 25–6, 28,
 32, 117, 123n9, 136, 159, 232, 252, 269,
 300–3, 300n5, 303n7, 305, 306, 310,
 379, 393, 396
 Smith, Vernon L. 14n5, 39, 52, 81, 95, 164,
 176, 178
 Sober, Elliott 124n10, 136, 137
 social atomization 322
 social cohesion 118, 127, 130, 281
 social cooperation 4, 18, 152, 153, 229, 231,
 233, 236
 in Austrian school 240–4
 social costs 201, 333, 562
 social dilemmas 128, 129, 149, 162
 social economics 5, 210, 211, 213
 understanding of positive-normative
 distinction 219–21
 vision: economy embedded in
 society 218–26
 strong thesis 220–1
 weak thesis 220, 221
 social ethics (Hayek) 138–55, 254n7, 256n6
 social evaluation 423, 424, 431, 434, 436,
 438, 439, 441
 social evolution 139, 139n1, 148, 150
 social hierarchy 128, 319
 social identity 183, 189
 and morality 196
 on lying 191
 social inequalities 264
 social instincts 126, 127
 social insurance 334
 social justice 96, 327, 374
 social media 273, 307, 386
 social morality 142n3, 143, 146, 147, 149, 154, 155
 social movements 191, 285
 social norms 101, 138, 140, 142, 142nn2,3, 149,
 163, 168, 175, 188, 193, 194, 251
 social organization 88, 144, 152
 social philosophy 155, 232, 239
 social policies 119, 213, 214, 219, 345, 374

- interventions-based 225–6
- prohibitions-based 224–5
- social progress 13, 15–18, 21, 25, 86
- social provisioning 262
- social psychology 39, 152
- social rationality 161, 191, 425, 427, 432, 435, 439
- social relations 119, 130, 210, 218, 241, 262, 350
- social responsibility 80, 201, 202, 286, 358
 - of finance 371–5
- social risk, normative economics of 423–41
- social rules 139, 142, 142n3, 145, 148, 152, 153, 155, 240
- social sciences 214, 219, 233, 477n2, 500, 545, 550
- social security 250, 265, 332
- social welfare 6, 54, 66, 359, 367, 534, 535, 585
 - functions *see* SWFs
- socialism 107, 144, 230, 231, 234, 236, 316, 317, 322, 324–32
- Socialist Calculation Debate 325
- socialization 131
- socially beneficial behavior 46
- socially responsible investment 371, 373–5
- society
 - economy embedded in 218–26
 - embedded in economy 213–18
- socio-economic systems 130
- sociological studies 281
- Socrates 39, 43
- soft budgets 333
- solidarity 6, 117, 124, 332, 499, 500, 517
 - fairness, and compassion 507–10, 511–13, 515, 516
- Sorell, Tom 366, 373, 379
- Soros, George 362, 379
- Soskice, David 331, 335
- South Korea 475
- Soviet Union 316, 330
- Spain 299
- specialization 15, 16, 161, 163, 170, 172, 233, 236, 476
- Spencer, Herbert 384–5, 396
- Spiegelhalter, David 448, 457, 473
- spirit of liberty 319, 322
- Spriggs, William E. 574, 579
- St Petersburg paradox 586
- state theory of money 360
- static efficiency 86
- Steadman, Lyle B. 125
- Steele, Claude M. 564, 579
- Stein, Michael A. 535, 542
- Steinbock, Bonnie 525, 542
- Stephenson, Matt 193, 206
- Sterelny, Kim 151n14, 159
- Steuart, James 252, 269
- Stewart, Dugald 232, 247
- Stigler, George 39, 52
- Stiglitz, Joseph E. 86, 95, 234, 247, 589, 604
- Stoics 4, 46, 47, 48, 48n7, 49
- Stone, Christopher 545n1, 558
- Storr, Virgil Henry xxi, 4, 13–29, 32, 33, 240, 270, 278, 281, 285, 289
- Stout, Lynn 341, 342, 357
- Stracca, Livio xv
- Strange, Susan 366, 379
- Streeck, Wolfgang 316, 332, 334, 336
- strikes 395, 507, 507n6
- Stringham, Edward 480, 495
- Strong Pareto 400, 403, 405n7, 410, 410n14, 411n19, 412
- strong reciprocity 128
- Strudler, Alan 370, 379
- stuffing 368
- Stuntz, William J. 146, 159
- subjectivism 231, 244
- subsistence approach 259, 260
- Sugden, Robert 46, 50, 139, 159, 502n3, 519
- Sullivan, Roger J. 56n4, 57n6, 75
- Sumner, L.W. 401, 422, 501, 520
- Sunstein, Cass R. xv, 308n10, 315, 521, 531n2, 533, 534, 541, 542
- Suojanen, Maria 439, 444
- supply 24, 191, 239, 263, 264, 298, 322, 338, 340, 344–8, 385, 416, 491, 507, 577
 - curve 345, 383
 - shifting down 354
 - decisions 344, 345, 354
 - and demand 226, 277, 344, 347, 382
- surplus value 253, 321
- surprise seeking 198–200
- survival 38, 44, 49, 121, 124–9, 162, 257, 262, 447, 455, 457, 461, 462, 549, 557
- sustainability 86, 264, 374, 552

sustainable economic growth 244
 sustaining economic activities 262–4
 Sutherland, Patrick 493
 Swanton, Christine 46, 53
 Swedberg, Richard 318, 336
 Sweden 201
 SWFs (social welfare functions) 399–420,
 478n3
 and cost-benefit analysis (CBA) 399–420,
 423–5, 427, 428, 430–4, 438–40,
 478, 521
 generalized utilitarian family of 400, 418
 prioritarian SWFs 400
 and tax system 416–17
 theory 404–10
 what type? 417–20
 Switzerland 507
 sympathy 1, 4, 18, 18n8, 19, 20n13, 41, 117, 123,
 124, 126, 127, 129, 172, 241, 299, 301,
 302, 401n4, 470, 481, 482
 circles of 22, 28
 systemic risk 487
 and financial crises 371–2
 Szech, Nora 201, 204

T

taboos 165
 tailgating 368
 Taleb, Nassim N. 487, 490, 495
 taming-the-market view 225–6
 Tancredi, Laurence 126n13, 136
 tariff(s) 108, 299
 penances 189
 tax evasion 186, 193, 198
 tax system 416–17
 taxation 107, 108, 235, 512, 514
 taxonomy of harmed or harmful
 conditions 484
 Taylor, Charles 277, 289
 Taylor, Fred M. 325
 Taylor, Paul C. 566, 579
 Taylor, Robert 392, 396
 Taylor-Robinson, David C. 448, 473
 team incentives 195
 technological progress 321
 technology 147, 321, 374, 391
 ethics of 97
 Teichgraeber, Richard F. 14n5, 33

Teichman, Doron 523, 543
 Temkin, Larry S. 418, 422, 600, 606
 Ten Commandments 79
 Thaler, Richard H. 308n10, 315, 502n3, 519
 Thamotheram, Raj 371, 379
 theft 55, 296, 301, 370
 Thiem, Annika 251, 269
 third-party punishment 183, 193
 third-person criterion 252
 Thiroux, Jacques P. 130, 137
 Thomas, Abulkader 365, 380
 Thomas, Charlotte C.S. 18n8, 20n13, 33
 Thomson, Judith Jarvis 55, 75, 522, 526, 542
 Thomson, William 423, 445
 Thorbjorn, Knudsen xviii, 135
 threshold deontology 7, 61, 61n15, 527
 thresholds 522, 526, 529, 530, 535–8
 Throsby, David 270, 276, 277, 289
 Tims, Dana 50, 53
 Tinbergen, Jan 277, 289
 Tintner, Gerhard 587, 606
 Tisak, Marie S. 128, 137
 Titmuss, Richard M. 201, 207, 284, 289
 Tocqueville, Alexis de 5, 318–20, 318n4, 322,
 323, 330, 335
 on capitalism and democracy 318–20
 Todorova, Zdravka 253, 269
 Tomasi, John 308n12, 315
 Tomaskovic-Devey, Donald 574, 579
 Tomlin, Ruth 271
 Torah 79
 tort law 70, 400n5, 482
 totalitarian socialism 329
 totalitarianism 327, 328
 Towse, Ruth 270, 289
 trade-tested progress 280
 trading networks 548, 549
 training 384, 392, 395, 568, 570
 transdisciplinary analyses 138–9
 transdisciplinary work 209
 Traub, Stefan 439, 445
 trauma 485
 Treich, Nicolas 433, 444
 Triple Theory (Parfit) 90–1
 Trivers, Robert L. 125, 126n12, 137
 trolley problem 55, 56n3, 526
 Tronto, Joan 250, 251, 269
 Trump, Donald J. 517

trust 29, 37, 41, 46, 47, 57, 81, 89, 161, 161n1,
164, 168–9, 173–6, 176n11, 184, 193,
196, 272, 282n11, 332, 368, 382, 506,
545, 550, 551
calculative 168
generation of 36
truth-telling 181, 182, 184, 185, 187, 190, 192,
194, 196, 198, 524
Tullock, Gordon 333, 336
Tuomala, Matti 399, 422
Turiel, Elliot 128, 137
Turnbull, Colin 142n6, 159
two-gender system 251
Tyler, Tom R. 122, 137
tyranny of the majority 320

U

Ubel, Peter A. 295n1, 307n9, 308, 315, 447,
455, 473
Uber 387
ul Haq, Mahbub 86
Ulen, Thomas 521, 533, 539, 542
Ulrich, Peter 249, 254, 255, 256, 257, 258, 269
ultimate ends 89–90, 109
underwriting money 361
unemployment 6, 54, 233, 235, 237, 332, 334,
382, 386, 389, 392, 479n4
benefit 274, 441
involuntary 382–5
rate 385
unifying across frameworks 90–1
unions 320, 390, 393–5
United Kingdom 232
United Nations 249
United States 28, 69, 86, 316, 328, 331, 334,
342, 360, 386, 395, 488, 488n10,
499, 503, 505, 511, 513, 559, 560, 563,
567, 574n6
Bureau of the Census 559n1, 567, 580
Constitution 59, 79
Federal Reserve 195
Supreme Court 85
universal earner model 260
Utikal, Verena 186, 207
utilitarianism 54, 59, 66–8, 72, 78, 80, 82, 90,
91, 123n8, 149, 418–20, 424–7, 430,
434, 435, 438–40, 446, 508, 523, 535,
569, 585, 595–7, 599, 601
vs. fairness 424
Kaldor-Hicks, and Pareto 66–7
utilitarian theorem (Harsanyi) 425
“utility,” do not use to denote a person’s
good 584–8
utility functions 13, 90, 138, 337, 400–3,
409, 410, 577, 585, 586, 588,
594, 595
and well-being and preferences 401–2
utility-maximization *see* utilitarianism

V

vaccination 108, 436, 437, 460, 499
vs. prevention 436
vaginal birth and C-section 449, 458, 459
Valentin, Andres 373, 376
value-based economics 91
van Staveren, Irene 3, 8
Vanberg, Viktor 120, 137
Vanderschraaf, Peter 151, 152, 159
Varian, Hal 384, 396
Vassel, Ron A. 524, 542
Vaughn, Karen 147, 159
veil-of-ignorance arguments 426, 438–40
Veblen, Thorstein B. 253, 269, 284, 289
Ven, Bert W. van de 369, 377
veneer theory of morality 126
Vermeule, Adrian 531n2, 542
vertical and horizontal pluralism 78–80
vice of markets 306–8
Villeval, Marie Clarie 203
Viner, Jacob 26, 33
violent revolutions 310
virtue 1–4, 18, 19n9, 21, 26, 27, 29, 61, 62, 64,
65, 78–81, 84, 87–91, 106, 122, 123,
145, 169, 182, 209, 270, 280, 281, 283,
286, 287, 304, 311, 312, 316, 317n2,
349n2, 355, 364, 484, 503, 504,
509, 532n4, 545, 555
and economics 34–50
ethics 18, 44, 59, 59n11, 79, 81, 209, 210
fitting virtue to ethical behavior 46–9
what is it? 38–40
virtuous actions vs. appropriate actions 47
Viscusi, W. Kip 535n6, 542, 567, 580
Visser, Hans 279, 289
Visser, Wayne A.M. 365, 380
von Neumann, John 587, 588, 606

von-Neumann Morgenstern (vNM)

utility functions 402

Voorhoeve, Alex 437, 441, 445, 458n4, 466,

466n8, 467, 471, 472, 473, 507n6,

520, 588, 589, 597, 598, 600, 605

objection to prioritarianism 467–8

voting 190, 191, 560, 562

W

Wade, Michael J. 124n10, 137

Waerness, Kari 250, 269

Wakefield, Jerome C. 500n2, 520

Waldfoegel, Joel 41, 53

Walras, Léon 117, 137, 320, 336

Walsh, Adrian 338, 357, 363, 380

Walton, Michael 270, 281, 282, 289

Walzer, Michael 270, 289, 393–4, 396

Wang, Shing-Yi 187, 206

war 196, 236, 298, 299, 304, 310, 332, 349, 564

Ward, Aidan 371, 379

Warde, Ibrahim 365, 366, 380

Warner, Michael 251, 269

Warren, Elizabeth 551, 552n7

Watzman, Asor 538

wealth 23, 24, 25, 28, 48, 48n7, 59, 68, 107, 108,

172, 238, 254, 256, 284, 302, 303,

308–11, 316, 318, 318n3, 319, 333, 358,

365, 372, 438, 483, 484, 507, 515, 552

of nations 21, 232

Weatherford, Jack 360, 380

Weber, Heloise 373, 380

Weber, Max 230, 235, 247, 274, 279, 281, 289,

364, 380

Wei, Maxine xv

Weil, David 389, 396

Weinstein, Jonathan 436, 441

Weisberg, Michael 139, 159

welfare 1, 4, 41, 54, 55, 63, 78, 81, 89–91, 99,

138, 139, 162, 163, 165, 167–9, 171,

174–6, 229–31, 257, 278, 280, 326,

329, 330, 331, 334, 337, 351n3, 355, 358,

359, 367, 537, 538, 560, 583–5, 587

assessment 86, 109

consequentialist evaluations of 82–7

economics 6, 89, 91, 98, 98n1, 109, 110, 230,

231, 235, 236, 399, 424, 426, 438, 439,

441, 475, 522–4, 527, 528, 530, 532,

533, 535, 583

Kantian implications for 65–72

neoclassical welfare theory 82–3

and risk and policy 397–493

state 193, 329–31, 425

welfarism 6, 109, 149, 399, 412, 420, 426,

479n4, 490, 535

welfarist consequentialism 399

well-being 1, 4, 6, 21, 26, 27, 47–9, 54, 57, 59,

66, 68, 72, 77, 85, 86, 89, 97–105,

108–11, 120, 181, 182, 221–4, 235, 259,

262, 263, 281, 319, 322, 333n14, 387,

393, 395, 399–402, 404–20, 433, 434,

437, 501, 502, 511, 512, 523, 527, 529,

532, 535, 569, 571, 584–99, 603

achievement 222

CBA is rough proxy for overall 414–15

and freedom 97

freedom 104

in developing countries 97

invariance to rescalings of well-being

measure 408

opportunities for 104

and preferences and utility

functions 401–2

types of comparisons 406

well-governed society (Smith) 301, 302,

312, 313

Werhane, Patricia H. 14, 33, 370, 380

Werlhof, Claudia von 260

wertfrei (value-free) social science 230

Wertheimer, Alan 384, 396, 490n12, 495

Wesley, John 364, 380

West, E.G. 26, 33

Westermarck, Edward 126n11, 137

Western, Bruce 570

Weymark, John A. 404, 421, 422, 426, 427,

438, 444, 445, 452, 452n3, 472, 473,

597, 606

Whately, Richard 89

Whigs 149, 153

whistle-blowers 193

White, Mark D. xv, xxi, 1–7, 54–73, 75, 76, 93,

227, 334, 471, 480, 483n8, 488, 493,

495, 524n1, 529, 532n4, 539, 543, 578

Whiten, Andrew 152, 158

Whiting, Jennifer 59n11, 73

Wicksteed, Philip H. 131, 137, 175, 178, 231,

240, 241, 243, 247

- Wiener, Jonathan B. 399, 422
 Wieser, Friedrich von 324
 Wiessner, Polly 128n15, 137
 Wight, Jonathan B. xxi, 3, 4, 8, 29n19, 33,
 77–93, 95, 524n1
 Wikler, Daniel 461n7, 472
 Wilber, Charles K. 3, 7
 Williams, Bernard 123n8, 136, 446n1, 473, 527,
 528, 532, 543, 551, 553, 558
 Williams, Evan 493
 Williams, George C. 124, 137
 Williamson, Oliver 168, 178
 willingness to pay (WTP) 403, 524
 Willott, Elizabeth 557, 558
 willpower 60–2
 Wilson, Bart J. 88, 95, 161n2, 164, 176,
 177, 178
 Wilson, David Sloan 124, 124n10, 136, 137,
 148, 150, 151, 153, 159
 Wilson, Edward O. 124n10, 137, 150n12,
 152, 159
 Winch, Donald 26, 33
 Wolf, Brianne 29, 33
 Woltz, Kaitlyn xxi, 5, 229–45
 women's rights 249, 256
 Wonnell, Christopher T. 531n2, 536, 543
 Woodmansee, Martha 270, 290
 Woodruff, Michelle 493
 work
 councils 394
 vs. labor 252
 workers 16, 25, 26, 80, 200, 249n2, 257, 258,
 283, 321, 331, 340, 355, 364, 381–6,
 388–95, 491, 531, 535, 563–5, 567, 568,
 570, 572, 573, 575–7
 workplace safety 334
 World Bank 85, 198
 Wozniak, Abigail 575, 580
 wrong, multiple approaches to understanding
 right from 79
 wrongfulness 68, 70, 72
 wrongs, rights, and externalities 67–71
 Wu, Alice H. 87, 95
 Wu, George 523, 541
 Wuerth, Julian 59n11, 74
- X**
 Xenophon 259
 Xu, Yongsheng 110, 113
- Y**
 Yair, Roi 538
 Yamagishi, Toshio 168, 178
 Yunus, Muhammad 372, 373, 380
- Z**
 Zaire 28
 Zak, Paul J. 81, 95, 126n13, 132, 137, 308n11, 315
 Zamir, Eyal xxi–xxii, 6, 55n2, 61n15, 65n21,
 520–38, 543
 Zanden, Jan Luiten Van 105, 113
 Zeckhauser, Richard 564, 579
 Zein-Elabdin, Eiman O. 259, 269
 Zelizer, Viviana A. 271, 284, 290
 Zeller, Adam 259
 Zerbe, Richard O., Jr., 70n28, 76, 533, 534, 543
 Ziliak, Stephen T. 91–2, 94
 Zuber, Stéphane 428, 430, 432, 440, 441, 442,
 443, 445
 Zuidhof, Peter W. 277, 278, 290