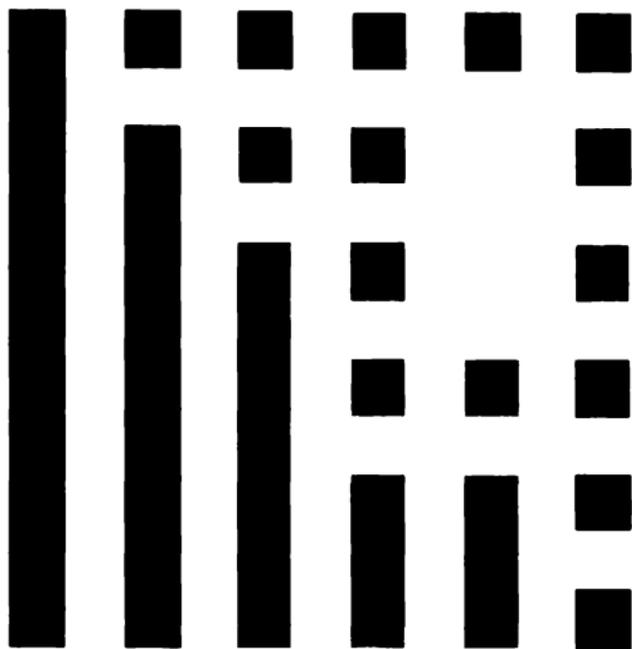


UTILITARIANISM AS A PUBLIC PHILOSOPHY



ROBERT E. GOODIN

Utilitarianism, the great reforming philosophy of the nineteenth century, has today acquired a reputation for being a crassly calculating, impersonal philosophy unfit to serve as a guide for moral conduct. Yet what may disqualify utilitarianism as a personal philosophy makes it an eminently suitable guide for public officials in the pursuit of their professional responsibilities.

Robert E. Goodin, a philosopher with many books on political theory, public policy, and applied ethics to his credit, defends utilitarianism against its critics and shows how it can be applied most effectively over a wide range of public policies. In discussions of such issues as paternalism, social welfare policy, international ethics, nuclear armaments, and international responses to the environmental crisis, he demonstrates what a flexible tool his brand of utilitarianism can be in confronting the dilemmas of public policy in the real world.

Written in a lucid, nontechnical style, these essays will interest a large cross section of the academic community concerned with, and teaching courses on, public policy, whether they be in departments of philosophy, political science, law, or economics.

Utilitarianism as a public philosophy

Cambridge Studies in Philosophy and Public Policy

GENERAL EDITOR: Douglas MacLean

The purpose of this series is to publish the most innovative and up-to-date research into the values and concepts that underlie major aspects of public policy. Hitherto most research in this field has been empirical. This series is primarily conceptual and normative; that is, it investigates the structure of arguments and the nature of values relevant to the formation, justification, and criticism of public policy. At the same time it is informed by empirical considerations, addressing specific issues, general policy concerns, and the methods of policy analysis and their applications.

The books in the series are inherently interdisciplinary and include anthologies as well as monographs. They are of particular interest to philosophers, political and social scientists, economists, policy analysts, and those involved in public administration and environmental policy.

Other books in the series

Mark Sagoff: *The Economy of the Earth*

Henry Shue (ed.): *Nuclear Deterrence and Moral Restraint*

Judith Lichtenberg (ed.): *Democracy and the Mass Media*

William A. Galston: *Liberal Purposes*

Elaine Draper: *Risky Business*

R. G. Frey and Christopher W. Morris: *Violence, Terrorism, and Justice*

Douglas Husak: *Drugs and Rights*

Ferdinand Schoeman: *Privacy and Social Freedom*

Dan Brock: *Life and Death*

Paul B. Thompson: *The Ethics of Trade and Aid*

Jeremy Waldron: *Liberal Rights*

Steven P. Lee: *Morality, Prudence, and Nuclear Weapons*

Bernard E. Rollin: *The Frankenstein Syndrome*

*Utilitarianism as
a public philosophy*

ROBERT E. GOODIN

RESEARCH SCHOOL OF SOCIAL SCIENCES
AUSTRALIAN NATIONAL UNIVERSITY



Published by the Press Syndicate of the University of Cambridge
The Pitt Building, Trumpington Street, Cambridge CB2 1RP
40 West 20th Street, New York, NY 10011-4211, USA
10 Stamford Road, Oakleigh, Melbourne 3166, Australia

© Cambridge University Press 1995

First published 1995

Printed in the United States of America

Library of Congress Cataloging-in-Publication Data

Goodin, Robert E.

Utilitarianism as a public philosophy / Robert E. Goodin.

p. cm.

Includes bibliographical references.

ISBN 0-521-46263-0 (hard). - ISBN 0-521-46806-X (pbk.)

1. Utilitarianism. 2. Political science - Philosophy. I. Title.

BB43.C66 1995

171'.5 - dc20

94-3385

CIP

A catalog record for this book is available from the British Library.

ISBN 0-521-46263-0 hardback

ISBN 0-521-46806-X paperback

For Brett and Ross

Contents

<i>Preface</i>	<i>page ix</i>
<i>Acknowledgments</i>	<i>xi</i>

PART I: INTRODUCTION; MORAL BASES OF STATE ACTION

1 Utilitarianism as a public philosophy	3
2 The state as a moral agent	28

PART II: MORALITY, PUBLIC AND PRIVATE

3 Do motives matter?	47
4 Government house utilitarianism	60

PART III: SHAPING PRIVATE CONDUCT

5 Responsibilities	81
6 Distributing credit and blame	88
7 Apportioning responsibilities	100

PART IV: SHAPING PUBLIC POLICIES

Respecting and overriding preferences

8 Liberalism and the best-judge principle	119
9 Laundering preferences	132
10 Heroic measures and false hopes	149
11 Theories of compensation	160

Contents

Ensuring social security

12	Stabilizing expectations	183
13	Compensation and redistribution	207
14	Basic income	228
15	Relative needs	244

International ethics

16	What is so special about our fellow countrymen?	265
17	Nuclear disarmament as a moral certainty	288
18	International ethics and the environmental crisis	306

	<i>References</i>	321
	<i>Name index</i>	347

Preface

This book constitutes my response to a challenge, laid by Peter Singer in the friendliest possible way over drinks one day in 1988, for me to write up my (by implication, eccentric) version of utilitarianism someday. I was slow in rising to the challenge, in part because I was relatively uninterested in purely metaethical discussions and supposed that I had little to contribute to them. My interest has always lain more in practical questions of what we should do, collectively more than individually and publicly more than privately. In the end, however, I have come to realize that that is in itself the kernel of a metaethical defense of utilitarianism – one that is both novel, at least in contemporary terms, and powerful, at least within its chosen scope.

If the idea of this as a book has been slow in coming, the individual chapters had been writing themselves for some time. All these chapters except the introduction were originally either freestanding articles or chapters, published in a disparate array of journals and collections over the best part of a decade. I am grateful to the editors and publishers of all those journals and books for permission to reproduce those materials here. I am especially grateful to Cambridge University Press for this opportunity to draw together these papers in a way that will highlight their important collective message. Those original essays have been lightly edited for this collection, to draw out their common themes while preserving something of their original flavor.

Many people have helped me work through these arguments over the years. Acknowledgments specific to particular chapters appear as a note at the end of each. I should record here my larger and recurring debt to a few people with whom I have discussed these issues repeatedly over many years. They include Brian Barry, Jon Elster, Debbie Fitzmaurice, Russell Hardin, Julian Le Grand, David Miller, Onora

Preface

O'Neill, Philip Pettit and Andy Reeve. Like my young sons' to whom this collection is dedicated, this book has been written around their frequent interventions and often in stubborn disregard of much of their best advice.

*Canberra
November 1994*

Acknowledgments

Chapter 2 reproduces material from "The State as a Moral Agent," in *The Good Polity: Essays in the Normative Theory of the State*, ed. Alan Hamlin and Philip Pettit (Oxford: Blackwell, 1989), pp. 123–39, by permission of the copyright holders, Blackwell Publishers Ltd.

Chapter 3 reproduces material from "Do Motives Matter?" *Canadian Journal of Philosophy*, 19 (September 1989), pp. 405–20, by permission of the copyright holders, the Canadian Journal of Philosophy.

Chapter 4 reproduces material from "Government House Utilitarianism," in *The Utilitarian Response*, ed. Lincoln Allison (London: Sage, 1990), pp. 140–60, by permission of the copyright holders, Sage Publications Ltd.

Chapter 5 reproduces material from "Responsibilities," *Philosophical Quarterly*, 36 (January 1986), pp. 50–6, by permission of the copyright holders, The Philosophical Quarterly.

Chapter 6 reproduces material from "Divorcing Liability from Responsibility," *Archiv für Rechts- und Sozialphilosophie*, 74 (1988), pp. 359–67, by permission of the copyright holders, Frans Steiner Verlag Wiesbaden GmbH.

Chapter 7 reproduces material from "Apportioning Responsibilities," *Law and Philosophy*, 6 (Fall 1987), pp. 167–85, © 1987 by D. Reidel Publishing Company. That material is reproduced here by permission of Kluwer Academic Publishers.

Chapter 8 reproduces material from "Liberalism and the Best-Judge Principle," *Political Studies*, 38 (June 1990), pp. 181–95, by permission of the copyright holders, Blackwell Publishers Ltd.

Chapter 9 reproduces material from "Laundering Preferences," in *Foundations of Social Choice Theory*, ed. Jon Elster and Aanund Hylland (Cambridge: Cambridge University Press, 1986), pp. 75–101, by permission of the copyright holders, Cambridge University Press.

Chapter 10 reproduces material from "Heroic Measures and False

Acknowledgments

Hopes," in *Philosophy and Medical Welfare* (Supplement to *Philosophy*, vol. 63, 1988), ed. Susan Mendus and Martin Bell (Cambridge: Cambridge University Press for the Royal Institute of Philosophy, 1988), pp. 17-32, by permission of the copyright holders, Cambridge University Press.

Chapter 11 reproduces material from "Theories of Compensation," *Oxford Journal of Legal Studies*, 9 (Spring 1989), pp. 56-75, by permission of the copyright holders, Oxford University Press.

Chapter 12 reproduces material from "Stabilizing Expectations: The Role of Earnings-related Benefits in Social Welfare Policy," *Ethics*, 100 (April 1990), pp. 530-53, © 1990 by The University of Chicago. That material is reproduced here by permission of the University of Chicago Press.

Chapter 13 reproduces material from "Compensation and Redistribution," in *Nomos XXXIII: Compensatory Justice*, ed. John W. Chapman (New York: New York University Press, 1991), pp. 143-77, by permission of the copyright holders, New York University Press.

Chapter 14 reproduces material from "Toward a Minimally Presumptuous Social Welfare Policy," in *Arguing for Basic Income*, ed. Philippe Van Parijs (London: Verso, 1992), pp. 195-214, © 1992 by Verso. That material is reproduced here by permission of the publishers.

Chapter 15 reproduces material from "Relative Needs," in *Needs and Welfare*, ed. Alan Ware and Robert E. Goodin (London: Sage, 1990), pp. 12-33, by permission of the copyright holders, Sage Publications Ltd.

Chapter 16 reproduces material from "What Is So Special about Our Fellow Countrymen?" *Ethics*, 98 (July 1988), pp. 663-86, © 1988 by The University of Chicago. That material is reproduced here by permission of the University of Chicago Press.

Chapter 17 reproduces material from "Nuclear Disarmament as a Moral Certainty," *Ethics*, 95 (April 1985), pp. 641-58, © 1985 by The University of Chicago. That material is reproduced here by permission of the University of Chicago Press.

Chapter 18 reproduces material from "International Ethics and the Environmental Crisis," *Ethics and International Affairs*, 4 (1990), pp. 91-105, © 1990 by the Carnegie Council on Ethics & International Affairs. That material is reproduced here by permission of Editor on behalf of the copyright holders.

PART I

**Introduction:
Moral bases of state action**

Utilitarianism as a public philosophy

Lecturing on an eminent Victorian at the height of the Vietnam War, Stuart Hampshire passes this telling judgment on the state of contemporary utilitarianism:

British utilitarianism was a school of moral thought . . . which set out to do good in the world, even though it was only a philosophy; and it . . . succeeded in large part over many years in this aim. [It] became part of the ordinary furniture of the minds of . . . enlightened persons, who would criticize institutions, not from the standpoint of one of the Christian churches, but from a secular point of view. . . . The utilitarian philosophy, before the First World War and for many years after it, . . . was still a bold, innovative, even a subversive doctrine, with a record of successful social criticism behind it. I believe that it is losing this role, and that it is now an obstruction.¹

Sharing both Hampshire's admiration for utilitarianism's earlier accomplishments and his dismay at many of its current tendencies, I aim in this book to rehabilitate a certain form of utilitarianism and to restore it to the critical social role it once so proudly boasted.

Utilitarianism is a doctrine that in its standard nineteenth-century formulation, directs us to produce "the greatest happiness."² In its most useful modern reformulation, it is "the moral theory that judges the goodness of outcomes – and therefore the rightness of actions insofar as they affect outcomes – by the degree to which they secure the greatest benefit to all concerned."³

¹ Hampshire [1972] 1978, p. 1.

² This is the phrase favored by Bentham ([1789] 1970, chap. 1, sec. 1, note) and his nineteenth-century friends and foes alike (Lively and Rees 1978). It supplanted Bentham's ([1776] 1988, p. 3) earlier, careless phrase – "the greatest happiness of the greatest number" – which critics always delight in describing as a mathematical impossibility (Hardin 1988, pp. 21–2).

³ Hardin 1988, p. xv.

Utilitarianism is an ethical theory with political consequences. It is an ethical theory, in the sense that it tells us what is right and wrong, good and bad. It is political, in that some of its most central pronouncements touch upon the conduct of public life. Indeed, it purports to provide a complete political theory, a complete normative guide for the conduct of public affairs.

An "ethic" is, strictly speaking, a theory of the good and bad, right and wrong quite generally. The term has, however, come primarily to connote more narrowly a theory of right conduct at the level of personal conduct. Ethics has come to be seen, quintessentially, as an answer to the question of "what should I do?" What is central to ethics thus understood is our intimate, individual affairs. What it is that is right for us to do jointly, in the conduct of our public lives, is seen to be basically derivative from that.

Of course this line of thought is quite right, in one sense. From most modern perspectives, if not from certain more ancient ones, The Politics always has to be parasitic upon The Ethics. Any political theory that purports to tell us what we should do (in more than a crassly prudential or pragmatic sense of "should") needs an ethical theory of some sort or another to provide its normative bite. What I shall here be disputing is whether that normative theory necessarily has to be parasitic upon – to be rooted in, to have its primary application to, to be tested first and foremost against its implications for – personal conduct.

The thesis of this book is that at least one normative theory, utilitarianism, can be a good normative guide to public affairs without its necessarily being the best practical guide to personal conduct. It is right there, too, after a fashion. But special circumstances confound the direct application of utilitarianism to personal affairs, and in such circumstances utilitarianism itself recommends that people's conduct be guided by more indirectly utilitarian mechanisms – obeying rules of conduct or developing traits of character, themselves chosen on utilitarian bases, rather than trying to apply the utilitarian calculus directly in each instance.⁴

There are special circumstances governing public life, too, however. Just as the special circumstances of private life are such as to drive us away from utilitarianism in any direct form, so too are the special circumstances of public life such as to drive us toward it. Those special circumstances make public life particularly conducive to the forthright application of utilitarian doctrine. Indeed, in my view, they make it almost indecent to apply any other.

⁴ Adams 1976. Hare 1981. R. Brandt 1988; 1992, pp. 263–89.

The larger argument of this book is that there are strong interrelationships between political theory and public policy, each having much to teach the other. Most of the cross-fertilization to be discussed will be substantive in character. At the outset, however, let us look to a more methodological plane.

One of the more interesting phenomena in public policy, with clear counterparts in political theory, is the phenomenon of "a solution in search of a problem."⁵ This reverses the ordinary order of rational decision-making, which (according to certain particularly narrow-minded canons of rationality) ought to start with a problem and proceed linearly toward a solution. In another way, though, it makes perfectly good sense to go about business the other way around. Often we strike upon nifty tricks or techniques, having no immediate use for them or having initially developed them for one purpose to which we now find them ill suited; and we simply store them away in our tool kit for use later. Like any good artisan, when presented with problems we look first to our existing tool kit to see if we have something readily to hand that is suited to the task, rather than fashioning a wholly new instrument for every task.

In the realm of public policy-making, the clearest cases of this phenomenon, as so many others, come in connection with weapons acquisition. Take the case of the cruise missile, for example. It was first seriously mooted within the American military establishment as an unarmed decoy to draw Soviet fire away from B-52 bombers. Later proposals offered it, in armed form, as the essential weaponry for a "stand-off bomber" that could destroy distant targets without penetrating enemy airspace. Later still, it was offered in land- and sea-based form as a ground-hugging missile that could evade enemy radar.⁶ Across its various incarnations, the problems differ but the solution – the cruise missile – remains the same.

Within the realm of political theory, the wavering fate of utilitarianism is not very different. In the megalomaniacal phrases of its founder and early advocates, of course, utilitarianism was touted as a universal panacea for whatever ethically ails us – for ethical problems large and small, simple and complex, personal and impersonal, public and private, individual and collective.⁷

⁵ Olsen 1972. March and Olsen 1976.

⁶ Levine 1977.

⁷ Bentham ([1776] 1988; [1789] 1970) and the elder J. Mill ([1823] 1992) were bad enough, but worse still in this respect were the lesser lights among their circle, so ably satirized by Dickens in *Hard Times* ([1854] 1969).

In a way, they were right to present their theory in this fashion. The value theory underlying utilitarianism is such that it should apply equally to all those realms. There are no grounds, within utilitarianism, for restricting the scope of its own application. Should utilitarianism turn out to constitute a credible theory only within some more restricted realm, that would prove to be a serious (and perhaps fatal) flaw, for utilitarianism provides no internal grounds for circumscribing its own scope in such ways.

Theories with such universal pretensions inevitably present their opponents with easy targets. A single counterexample defeats a universal claim, and antiutilitarians have been most ingenious in concocting painfully cute counterexamples to embarrass utilitarians. The counterexamples have depicted merely (often barely) possible worlds, more often than probable ones; they have been contrived, more often than commonplace. But no matter. Its universalist pretensions make utilitarianism absolutely fair game for purveyors of such fantasies.

One perfectly viable alternative for utilitarians confronted with such fantastic scenarios is for them simply to wear any such embarrassment. What those counterexamples do – all that they do – is to conjure up a situation in which doing the utility-maximizing thing would lead to intuitively unappealing results. The circumstances they depict, however, are very far from those to which our standard intuitions are standardly shaped. (They involve things like promises to dying friends on otherwise unpopulated desert islands and “super efficient pleasure machines” and such like.)⁸ Precisely because of that, we may well decide that it is our intuitions rather than the prescriptions of our utilitarian moral theory that ought to be readjusted in such unusual circumstances.⁹ The right way to treat antiutilitarians who offer such contrived counterexamples might well be to “out-smart” them, embracing what they offer as a *reductio* as being actually the correct (however counterintuitive) solution to such (decidedly nonstandard) situations.¹⁰

Patently crazy counterexamples, however, only ever formed part of the critics’ case against utilitarianism. Another important strand has to do with the inappropriateness, in all sorts of ways, of utilitarianism as a code of personal conduct. These criticisms, too, are licensed by utilitarianism’s universalist pretensions. It has – given its universalist pretensions, it has to have – direct implications for personal conduct. So to some extent utilitarianism was always going to have to lay itself open to this line of attack.

⁸ Ross 1930, p. 39. Friedman 1947.

⁹ Goodin 1982b, pp. 8–12. Hardin 1988, pp. 22–9.

¹⁰ The technique, described in Dennett and Lambert’s (1978, p. 8) *Philosophical Lexicon*, is named after the distinguished utilitarian, J. J. C. Smart.

Friends of utilitarianism have rather tended to lead with their chins in this regard, though. There was an important shift among utilitarian writers that came somewhere between Sidgwick's 1874 *Methods of Ethics* (where public affairs loomed large) and G. E. Moore's 1903 *Principia Ethics* (where the greatest good is defined in terms of more private ideals, such as friendship and aesthetic appreciation).¹¹ Whatever the cause of that shift, its consequences could not have been more deleterious to the proper defense of the utilitarian cause.

Throughout the twentieth century, defenders of utilitarianism have been primarily concerned to defend it in its least plausible form, as a code of personal conduct. They have retained that focus even when, of late, reapplying it to public affairs.¹² The question of war is seen, by them, as one of whether we individually should kill enemy soldiers or defenseless civilians, rather than as one of whether we collectively should wage pointless wars.¹³ The question of famine is seen as one of whether we individually should send food to the starving, rather than as one of whether we collectively should work to reform social structures and consequent exchange entitlements.¹⁴

Attacks upon utilitarianism as a personal code are on powerful grounds. No one wants to run one's life like *Gradgrind*, the Dickensian parody of a good utilitarian.¹⁵ Furthermore no one can. The calculative load imposed by utilitarian maximization would absorb all one's time and attention, leaving none for actually acting on the conclusions of the calculations. In personal life, most dramatically, there simply has to be more scope for considerations of uncalculating affection, standing rules of conduct and qualities of character.

All of that is, if not exactly foreign to utilitarianism, at least a fair distance from its central precepts. With enough twisting and turning, of course, you can get from anywhere to anywhere.¹⁶ But that is no defense – the sheer fact that twisting and turning is required in itself constitutes a telling criticism.¹⁷ If that is where you want to end up – if what centrally matters, in an assessment of an ethical theory, is its ability to give a clear and coherent account of ordinary intuitions about how we ought to conduct our personal affairs – then it seems utilitarianism is not where you would most naturally want to start.

¹¹ Sidgwick [1874] 1907. Moore 1903.

¹² This is basic tenor even of such exemplary utilitarian texts as Singer (1979) and R. Brandt (1979; 1992, esp. chaps. 12–19).

¹³ Contrast *Philosophy and Public Affairs* treatments of this topic (Held, Morgenbesser, and Nagel 1974; M. Cohen, Nagel and Scanlon 1974) with Hare's (1957).

¹⁴ Contrast Singer 1972 with Drèze and Sen 1989.

¹⁵ Dickens [1854] 1969.

¹⁶ There are even good utilitarian grounds for twisting and turning, here: see Hardin 1988.

¹⁷ Barry [1979a] 1989a, p. 341.

Of course, utilitarianism does purport to give an account of how to lead our personal lives as well as our public ones. Furthermore, early, straightforward and frankly simpleminded applications of the utilitarian maxim to personal ethical dilemmas have, of late, given way to more sophisticated accounts taking due notice of the "limits of reason" facing isolated individual choosers in their own daily lives. The upshot is a utilitarian theory of personal morality much nearer ordinary intuition than we have formerly been led to expect.¹⁸

But that is a defense of utilitarianism at its weakest. If well defended on that front, utilitarianism might avoid being dismissed on those grounds alone. However good that defense, though, it is in the nature of defensive maneuvers that they only ever avert defeat. They never actually secure victory. The most we might hope is that those defenses of utilitarianism as a personal ethical code might prevent people from dismissing it. No one is going to embrace utilitarianism for the brilliance of its answer to those conundra alone.

II

The strength of utilitarianism, the problem to which it is a truly compelling solution, is as a guide to public rather than private conduct. There, virtually all its vices – all the things that make us wince in recommending it as a code of personal morality – loom instead as considerable virtues.

Consider first the raft of criticisms couched in terms of the *impersonality* of utilitarianism. Like all universalist philosophies, utilitarianism asks us to take "the view from nowhere."¹⁹ There is no obvious place within utilitarian theories for people's idiosyncratic perspectives, histories, attachments, loyalties or personal commitments.

That rings untrue to certain essential qualities of personal life. The essence of the communitarian challenge is that everyone comes from somewhere. There are no free-floating individuals, of the sort with which liberals generally, and utilitarians paradigmatically, populate their moral theories.²⁰ People have, and upon reflection we think they should have, principled commitments and personal attachments of various sorts.²¹

¹⁸ Hardin 1988.

¹⁹ Nagel 1986. Similarly agent-centered permissions or obligations: B. Williams 1973a; Scheffler 1982; Sen 1982b.

²⁰ This communitarian critique (Sandel 1982; 1984; cf. Kymlicka 1989 and Avineri and de-Shalit 1992) is echoed by many contemporary feminists (Mansbridge and Okin 1993).

²¹ Those who do not are said to be "rational fools," in Sen's term (1977a), or to lack "integrity," in B. Williams' (1973a). All of this, most effectively developed

As an account of the peculiar role responsibilities of public officials (and, by extension, of ordinary individuals in their public capacities as citizens) that vice becomes a virtue, though. Those agents, too, have to come from somewhere, bringing with them a whole raft of baggage of personal attachments, commitments, principles and prejudices. In their public capacities, however, we think it only right and proper that they should stow that baggage as best they can.

Complete neutrality might be an impossible ideal. That is another matter.²² But it seems indisputable that that is an ideal which people in their public capacities should strive to realize as best they are able. That is part (indeed, a central part) of what it is to be a public official at all. It is the essence of public service as such that public servants should serve the public at large. Public servants must not play favorites.

Or consider, again, criticisms revolving around the theme that utilitarianism is a coldly *calculating* doctrine.²³ In personal affairs that is an unattractive feature. There, we would like to suppose that certain sorts of actions proceed immediately from the heart, without much reflection much less any real calculation of consequences. Among intimates it would be extremely hurtful to think of every kind gesture as being contrived to produce some particular effect.

The case of public officials is, once again, precisely the opposite. There, it is the height of irresponsibility to proceed careless of the consequences. Public officials are, above all else, obliged to take care: not to go off half cocked, not to let their hearts rule their heads. In Hare's telling example, the very worst thing that might be said of the Suez misadventure was not that the British and French did some perfectly awful things (which is true, too) but that they did so utterly unthinkingly.²⁴

Related to the critique of utilitarianism as a calculating doctrine is the critique of utilitarianism as a *consequentialist* doctrine. According to utilitarianism, the effects of an action are everything. There are no actions which are, in and of themselves, morally right or wrong, good or bad. The only things that are good or bad are the effects that actions produce.²⁵

That proposition runs counter to certain ethical intuitions which, at

recently in Fletcher 1993, merely picks up themes first introduced in John Stuart Mill's ([1838] 1962, pp. 101-2) own "Essay on Bentham."

²² Goodin and Reeve 1989.

²³ In Lukes's (1993, p. 428) dystopia, Utilitaria, "Calculating is the notional obsession," which goes far toward explaining what he finds so awful about the place. See similarly Hampshire [1972] 1978, esp. pp. 5ff.

²⁴ Hare 1957.

²⁵ Pettit 1991.

least in certain quarters, are rooted deeply. Those who harbor a Ten Commandments view of the nature of morality see a moral code as being essentially a list of "thou shalt" and "thou shalt not" – a list of things that are right or wrong in and of themselves, quite regardless of any consequences that might come from doing them.²⁶

That may or may not be a good way to run one's private affairs.²⁷ Even those who think it is, however, tend to concede that it is no way to run public affairs. It is in the nature of public officials' responsibilities that they are morally obliged to "dirty their hands" – make hard choices, do things that are wrong (or would ordinarily be wrong, or would be wrong for ordinary private individuals) in the service of some greater public good.²⁸ It would be simply irresponsible of public officials (in any broadly secular society, at least) to adhere mindlessly to moral precepts read off some sacred list, literally "whatever the consequences."²⁹ Doing right though the heavens may fall is not (nowadays, anyway) a particularly attractive posture for public officials to adopt.

Yet other critics berate utilitarianism not so much for dictating maximization and the coldly calculating, consequentialistic attitude that that engenders as for what it would have us maximize. There is, such critics say, something necessarily *crass* about whatever utilitarians take as their maximand.³⁰ That maximand has subtly shifted over the years. Modern utilitarians now go well beyond Bentham's simple calculus of pains and pleasures, so it is no longer true to say that the utilitarianism is necessarily a crassly "hedonic" philosophy. But dress up their maximand as they will, modern utilitarians – to deserve the name at all – must necessarily be involved in maximizing satisfactions, somehow construed, of people, somehow specified.

²⁶ Anscombe 1958. B. Williams 1973a. Fried 1978.

²⁷ Whether it is depends essentially upon finding theological foundations for the commandments: "[I]f we take God and the soul out of the picture, the emphasis upon 'integrity' becomes a form of narcissism" (Barry [1979a] 1989a, p. 340).

²⁸ As is Weber's (1919) theme in "Politics as a vocation"; for modern restatements, see Walzer 1973, Nagel 1978 and Thompson 1987.

²⁹ Various writers try to find a halfway house in notions of the "unthinkable" (B. Williams 1973a, pp. 90–3) or the "morally impossible" (Hampshire [1972] 1978, pp. 9ff.) or the intolerably "cruel" (Shklar 1984, chap. 1) or "human rights" marking the limits of what politicians should even consider doing. But of course refusing to think about grossly evil outcomes is not in general a very good way of guaranteeing that they (or something worse) do not happen.

³⁰ Again, a central feature of Lukes's (1993, p. 428) caricature of his dystopia is that "Utilitarians are a distinctly phillistine people, who are disinclined to see utility in High Culture and never tire of citing the proverb that 'pushpin is as good as poetry,'" although Lukes does go on to acknowledge that "there is a minority tradition of trying to enrich the idea of 'utility' to include the more imaginative sides of life."

Genuinely "higher" concerns are genuinely hard to accommodate within utilitarianism. The higher they are – the more distant they are from anything connected (directly or indirectly, immediately or eventually or counterfactually) to actual or possible satisfactions of actual or possible people – the less of a place such higher ideals would seem to have within any genuinely utilitarian framework.

Aesthetes would bemoan the poverty of a life lived according to such purely utilitarian precepts. If (as utilitarians assert) beauty lies only in the eye of the beholder, then it amounts to no more than an illusion which we create for ourselves; and beauty is thus emptied of any independent value. If (as utilitarians assert) there is nothing that is good in and of itself, then there is no external standard by which to validate any of our subjective valuations. There is nothing to live for, nothing to die for. We are ourselves the only ultimate source of value in the universe. The external world of values is thus demystified, but at the same time it is also somehow diminished for us. None of us, aesthetes scold utilitarian opponents, would want to live our individual lives in a universe so empty of external meanings and values.

Here again, however, what would diminish private life is perfectly suitable for circumscribing public life. The classically utilitarian habit of asking "of what use is it to me?" (or in more public-spirited fashion, "of what use is it to us?") may be a crass way of judging what to do in one's personal affairs. There may well be some things that it would be good or bad, right or wrong for private individuals to be or to represent or to do, whether or not they would be of any use to anyone. Whatever we might say of the merits of that as a personal code, however, it has distinctly limited appeal as a guide for public officials.

Persisting in costly practices that do no one any conceivable good whatsoever is, as a public practice, simply perverse. Even if it would be right for private individuals uniformly to adopt a code of pointless self-sacrifice, it seems transparently wrong for public officials to impose such sacrifices upon any who refuse to undertake them voluntarily.¹¹ The root intuition here is perhaps best captured by saying that rulers have no right to wage holy wars – anyway, not ones waged on behalf of gods in whom their subjects no longer have any faith.¹²

III

Perhaps it is novel nowadays to look at utilitarianism as essentially a public philosophy. If so, the novelty is itself wholly new. In earlier

¹¹ Bentham [1789] 1970, chap. 2.

¹² Doing so in the service of shared ideals and common gods is another thing altogether, of course: that just amounts to organizing collective activities that people will, from their own personal perspectives, find satisfying.

times it was much more of a commonplace to suggest that utilitarianism constitutes a solution to public rather than personal moral problems, to defend it as a public philosophy rather than as a personal moral code.

That much is clearly suggested by reflection upon the corpus and the personal history of the founders of the utilitarian tradition.¹¹ Jeremy Bentham was famous in his own day primarily as a reformer of legal systems; James Mill as an essayist on government; John Stuart Mill as an essayist, social reformer and parliamentarian; John Austin as a jurist. The bulk of Bentham's large corpus is given over to constitutional codes, systems for penal reform, and such like. The two Mills fixed their focus equally firmly on public affairs ranging from histories of British India to political economy and women's suffrage.

Lest all that seem to amount to a post hoc reconstruction of what the nineteenth-century utilitarians were up to, consider the younger Mill's self-conscious reflection upon the utilitarian tradition which he inherited. In his memorial essay on Bentham, John Stuart Mill writes,

It is fortunate for the world that Bentham's taste lay rather in the direction of jurisprudential than of properly ethical inquiry. Nothing expressly of the latter kind has been published under his name, except the "Deontology" – a book scarcely ever . . . alluded to by any admirer of Bentham without deep regret that it ever saw the light.

Turning from what Bentham did badly to what he did well, Mill continues,

If Bentham's theory of life can do so little for the individual, what can it do for society? It will enable a society . . . to prescribe the rules by which it may protect its material interests. It will do nothing . . . for the spiritual interests of society . . . [W]hat a philosophy like Bentham's can do [is to] teach the means of organizing and regulating the merely business part of the social arrangements.¹²

That is a fair assessment of Bentham, and of the tradition to which he gave rise. And although the younger Mill himself aspired to do better, it is in the end fair reflection, too, of John Stuart Mill's utilitarian accomplishments.

IV

Over the years, there have grown up a great many diverse forms of utilitarianism.¹³ To some extent, this amounts to distinction-

¹¹ Halv y 1928. Plamenatz 1958.

¹² J. S. Mill [1838] 1962, pp. 104–6.

¹³ For a summary, see Smart 1967; cf. Lukes's (1993, pp. 427–8) more recent, breezier version. Much of what follows draws on Goodin 1991a, b.

mongering for the sheer sake of it. To some extent, the distinctions do matter – but mainly for the fine-grained forms of utilitarianism which might properly characterize personal moral codes but which cannot possibly characterize public ones. If my survey of a hundred years' worth of analytic work in this area seems cavalier, riding roughshod over distinctions which have been honed by many hands with loving care, that is only because those fine distinctions are irrelevant or worse for the rough-and-ready form of utilitarianism which is the only form that public officials can deploy (which, I would also say, is the principal form that philosophers have any business recommending; but that is another story, already told).

One dimension, already alluded to, along which these various utilitarianisms differ concerns the *content* of the utilitarian maximand. In its earliest Benthamite formulation, it was some state of mind ("happiness," "the balance of pleasure over pain") that was supposed to be maximized. That variant can be called "hedonic (or hedonistic) utilitarianism." Subsequent utilitarians, whilst preserving the basic spirit of that philosophy, watered down its hedonism. First came the recognition that people sometimes derive satisfaction from things that do not literally give them pleasurable buzz. That recognition gave rise to a revised utilitarianism that would have us maximize "preference satisfaction" (call this "preference utilitarianism"). Next came the recognition that people sometimes would derive satisfaction from something in a way that they do not presently recognize. That gave rise to a further revised utilitarianism that would have us maximize people's "interests" or "welfare" (call this "welfare utilitarianism").*

As among these options, the last is the form of utilitarianism that I shall be advocating in this book. That choice is an easy one. In this list, each successive refinement subsumes the former. Preference utilitarianism subsumes hedonic utilitarianism, if one assumes people ordinarily actually prefer (among other things) to experience pleasure and avoid pain. Welfare utilitarianism subsumes preference utilitarianism, if one assumes that promoting people's welfare interests will ordinarily lead to higher levels of preference satisfaction in the future.

* "Interests" or "welfare," insofar as they are (as here) to be distinguished from preference satisfaction, are best analyzed as references to resources or basic capabilities which will prove useful to people whatever their ultimate ends (Barry 1964; Sen 1985a). There is sometimes discussion of a fourth class of "ideal utilitarianism," the central claim of which is that certain ideals (things, attributes) are good independently of people's desire for them. That counts as a form of utilitarianism at all, I would argue, only if it can be assimilated under this last category. J. S. Mill's (1863) "higher pleasures" are arguably higher in precisely the sense of having a higher utility yield, if achieved. Arguably it is precisely that to which people are testifying when Mill would have us ask anyone who has experienced both higher and lower which they would prefer.

What is added with each refinement in the list is something which anyone working within the spirit of utilitarianism should surely want to incorporate among sources of utility (nonhedonic preferences, hitherto unacknowledged sources of satisfactions).

Although that choice of welfare utilitarianism is driven primarily by philosophical considerations, that choice is ratified by pragmatic considerations having to do with the peculiar circumstances of policy choice in the public sphere. People's pleasures are varied, their preferences idiosyncratic. People's basic interests, in contrast, are (at least at some suitable level of generality) pretty standard across all individuals (at least all individuals in any given society at any given moment). Policy makers are of necessity making choices that must affect a wide range of people at one and the same time. They might well be overwhelmed by the diversity of pleasures and preferences; and if that were the focus dictated by the most philosophically credible form of utilitarianism, they might be forced to abandon utilitarianism as a public philosophy altogether. Their task is rendered far more manageable by the relative commonality of people's basic welfare interests, which should instead be central to policy-makers' concerns when implementing what is in any case the most philosophically credible form of utilitarianism.

Different utilitarianisms also differ importantly in the *modal status*, as well as in the content, of the maximand that they specify. The crux of this issue is whether we should be concerned only with what is actually there – only with real people, their preferences, their pleasures and pains, their welfare – or whether we should allow our maximizing to range across all possible people (and their preferences, pleasures/pains, welfare). "Welfare utilitarianism" is my own favored variety, but the differences here in view come out most clearly with reference to "preference utilitarianism," so let us for the moment talk in those terms.

One strand of utilitarianism would have us maximize preference satisfaction across all possible preferences and all possible people. That, arguably, is the classically Benthamite approach. But it leads to "repugnant conclusions" for population policy. Maximizing utility in that way would lead to a population explosion, as we bring more and more people into the world up to the point where any new person's utility from being born is outweighed by the disutility that extra person's existence causes to everyone already there.¹⁷ That is what is wrong with maximizing preference satisfaction (or indeed anything else) across all possible people. What is wrong with maximizing preference satisfaction across all possible preferences is that it leads to

¹⁷ Parfit 1984, pp. 381–91.

absurdly “adaptive preferences.” If you cannot get what you want you should simply revise your preferences so you will want what you can easily get. The ancient Stoics thought that a good idea, but surely few of us would find the satisfaction of preferences chosen on that basis alone all that satisfying.³⁸

Another form of utilitarianism would have us focus instead on satisfying actual preferences of actual people. That would certainly avoid each of those problems just posed. But it would also debar utilitarians from doing one of the things that utilitarians have always most wanted to do. That is to use utilitarian criteria to guide policy-makers in choosing among a whole range of alternative devices (ranging from legal institutions to educational curricula) by which people are socialized in one way rather than another and, in so doing, by which their preferences are shaped in one way rather than another. Perhaps it is that aspiration itself that utilitarians ought to abandon. But on reflection perhaps there is no need for them to do so.

The example I deployed against letting utilitarianism range over all possible preferences is an argument couched in terms of what might be called “directly” adaptive preferences. Perhaps the satisfaction of easily satisfied preferences is unsatisfying when you have yourself intentionally chosen those preferences precisely because they will be easy to satisfy. But perhaps the satisfaction of such preferences would be not at all unsatisfying if those preferences had been inculcated in you by some external agency, some socialization device, which was itself intentionally chosen for just those attributes by some utilitarian social planner. Arm’s length, indirectly adaptive preference formation – which is, after all, the principal form that public policy-makers practice – might well be able to serve utility-maximizing ends, even if directly adaptive preferences cannot.³⁹

I am inclined to commend a form of utilitarianism which takes actual people as given, but which maximizes across all their possible preferences.⁴⁰ Perhaps the easiest way of making that case is merely

³⁸ Elster 1982. Barry (1989b) conjures up the telling example of how unsatisfying it would be to see yourself as a “fair weather fan” of a sporting team. If your team is a pack of hopeless losers, simply shift loyalties so you end up backing the winners instead. One cannot imagine that such people get much satisfaction out of seeing “their” team (a team that is “theirs” in such an attenuated sense) actually winning the championship.

³⁹ There are many goals which cannot be attained by agents aiming at them intentionally, but which can only be achieved if pursued indirectly. This is Elster’s (1983, pp. 43–108) important category of “essential by-products,” like sleep and spontaneity.

⁴⁰ That is not, of course, to say that any course of action is permissible just so long as it satisfies some preference that the person might possibly come to have – that is to miss the force of the phrase “maximizing [utility] across” all possible pref-

to recall that people's preferences are, in any case, bound to undergo changes between the time their choices are made and the time the consequences of those choices are felt. Clearly, in such circumstances, anyone moved by the spirit underlying utilitarianism should say that there we ought to cue on the preferences which those people will actually have, come the time the consequences are felt, rather than on the preferences which they had when the choice was made but which have long since disappeared.⁴¹

That form of utilitarianism has the further virtues of being consistent with our utilitarian desire to employ our principles to shape socialization and with the form of welfare-utilitarianism that we have other reasons for favoring. The most plausible way to construe "welfare interests" is, after all, as a form of "possible preferences" – as those things which people counterfactually would favor, under certain privileged circumstances which do not necessarily actually prevail.

Finally, to complete this short survey, various utilitarianisms differ according to what they regard as the appropriate *object* of the utility calculus. What is it that the utility calculus is to be used to choose? "Act utilitarianism" (which the received wisdom regards, I think wrongly, as the classical approach)⁴² says that each of our actions, one by one, should be chosen so as to maximize overall utility. "Rule utilitarianism," in contrast, says that we should employ the utility calculus to choose among rules (habits, norms, patterns of behavior) in such a way as to maximize overall utility.⁴³ That pretty well exhausts the traditional catalog of options.⁴⁴ To that traditional catalog have recently been added proposals for using the utility calculus to choose motivational structures, or character traits, which will in turn maximize overall utility.⁴⁵

ferences. That involves us, *inter alia*, in choosing that course of action which maximizes utility, given which of the possible preferences a person will most probably actually have at the time the effects of the actions are felt, or in inculcating preferences in people in such a way that the maximally satisfying possible preferences become people's actual ones.

⁴¹ Sen 1957. Goodin 1982b, pp. 39–56.

⁴² Certainly it is Sidgwick's ([1874] 1907). But Bentham ([1789] 1970) was inclined to use the utility calculus, more ambiguously, as a guide to both morals and legislation. Insofar as he used the utility calculus to guide our choice of legislation, constitutional codes, penal regimes and so on – in short, to choose systems of rules as well as discrete actions with a view to maximizing social utility – he might reasonably be classed among rule as well as act utilitarians.

⁴³ That, anyway, is the most defensible version of rule utilitarianism. A more extreme classical form holds that you should choose that act which would be required by a rule whose adoption would maximize utility (which may be different, if the utility-maximizing rule is not in place and generally obeyed).

⁴⁴ Lyons 1965. Smart 1967; 1973. Hare 1981. Hardin 1988.

⁴⁵ Adams 1976. R. Brandt 1988; 1992, pp. 263–89.

The crucial argument standardly offered in favor of act utilitarianism, as against any other form, goes as follows. In any given instance, either act utilitarianism and rule/motive utilitarianism recommend we do the same thing; or else act utilitarianism recommends that we do one thing, whereas following rules or motives chosen on utilitarian grounds would have us doing something else. In the former case, the other forms of utilitarianism are extensionally equivalent to act utilitarianism. Wherever they differ in their recommendations, they are inferior to act utilitarianism on straightforwardly utilitarian grounds. Slavishly following a rule or disposition chosen on utilitarian grounds, when more utility could be achieved by doing something else, cannot be the utilitarian thing to do.⁴⁶

That argument is compelling in its own terms. But its terms are not of this world, for individual moral agents or (more especially) for social policy-makers, either. The argument presupposes that we are able to perform utility calculations that typically range across an enormous number of individuals and options – and that we are able to do so reliably, instantaneously and costlessly. This is to assume away the “limits of reason” which characterize the real world for individual agents and, all the more so, for social policy-makers alike.⁴⁷

The arguments in favor of some form of rule utilitarianism at the social level (and, correlatively, some form of motive utilitarianism at the individual level) are built in large part out of those limits of reason. At the individual level, rules serve to maximize utility in the real world, in ways that act-by-act calculations of utility cannot, by being easier to communicate, easier to inculcate, easier to remember and easier to apply.⁴⁸

Most important, at the social level, rules are publicly accessible in a way that (given those crucial facts about the real world) private utility calculations are not. In an ideal world in which everyone had perfect information, anyone could replicate anyone else’s utility calculations. But in the real world, no one has perfect information; and, furthermore, each of us is privy to different bits of the overall picture.⁴⁹ Thus, what I would think is the best act-utilitarian action for you to perform, given what I know, may not be (typically will not be) the action that you, given what you know, think is the one that act-utilitarianism would most highly recommend to you.

That matters because much of the utility that utilitarians would

⁴⁶ Lyons 1965. Smart 1967; 1973.

⁴⁷ “Any argument that turns on perfect information, perfect calculation, and perfect theory is a house of cards, [and therefore] is almost entirely beside the point for a practical morality” (Hardin 1988, p. 17).

⁴⁸ Hare 1981. Hardin 1988.

⁴⁹ Hayek 1945.

have us maximize comes from coordinating the actions of a great many individual agents. Often the only way to maximize the utility that arises from my act is by knowing (or guessing) what others are likely to do. But knowing that with any certainty is for the reasons just given impossible (or impossibly costly) in a world populated by act-utilitarian agents. The best way to coordinate our actions with those of others, and thereby maximize the utility from each of our actions as individuals as well as from all of our actions collectively, is to promulgate rules (themselves chosen with an eye to maximizing utility, of course) and to adhere to them.⁵⁰

If adherence to rules is justified in these terms, then there is no risk of the sort of "rule fetishism" of which act utilitarians complain in their critiques of rule utilitarianism. We choose the rule we do because it is the one which, if generally followed, would yield more utility overall than would general adherence to any other rule. We follow that rule, in its various edicts, because we think that the effects of our doing so will maximize utility in most applications, either directly or indirectly (given the benefits of coordinating our actions with those of others, who can best guess what we will do by assuming us to be followers of known rules). We follow that rule in some applications, even where we are pretty confident that doing so will not maximize utility in that instance, because of the larger general benefits to be had by society at large (and, incidentally, by us individually) from our being known by others to be rule-followers. But when, just occasionally, following a rule would have truly grievous utility consequences, we would be perfectly well licensed to abandon the rule by the self-same logic that led us to adopt rules and generally to follow them. Rule utilitarians can, thus, lie to Nazis about the Jews hidden in their attic, in a way that Kantian rule-followers might find hard.⁵¹

V

In like fashion, the public rather than private application of utilitarian precepts helps us evade some of the most standard practical and practicality objections to the doctrine. Primary among them is the objection that utilitarianism can never be implemented in practice, because it

⁵⁰ All the standard arguments for rule utilitarianism seem to turn, at crucial points, on some such proposition (Hodgson 1967; Regan 1980; Hare 1981; Hardin 1988; Honoré 1993).

⁵¹ Hare (1981), e.g., is clear on this point. Some would say that this just amounts to act-utilitarianism as applied to the choice and application of rules. I nonetheless persist in describing this position as a kind of rule-utilitarianism in order to signal the important role in my analysis of social rules, even if they are perhaps merely act-utilitarian in their ultimate justification.

requires us to engage in impossible "interpersonal utility comparisons."⁵²

Whatever course of conduct we pursue, there are likely to be some people who gain and others who lose. If we are supposed to be maximizing utility across all those people, we obviously need some standard for comparing the gains of the one group with the losses of the other.⁵³ Critics, however, have long complained that we have no such way of getting inside another's head, experiencing the other's pleasures and pains as the other does.⁵⁴ In Jevons' famous phrase, "Every mind is inscrutable to every other mind [so] no common denominator of feeling is possible." That is said to render utilitarian maximization impossible, for all practical purposes.⁵⁵

In the absence of interpersonal comparisons, the most we can have are the weak Pareto-style comparisons upon which so much of modern welfare economics rest. We can say a state of affairs is indisputably superior to another if someone is better off and no one worse off. But that Pareto principle suffers from various defects. It is indecisive, leaving altogether too many alternatives unranked (in any large-scale application most options will make at least a few people worse off); and it is essentially conservative (redistribution of resources can never be justified on utility-maximizing grounds if we have no way to say that gainers would gain more than losers would lose). Such problems have driven welfare economists to a variety of ruses to reinstate something akin to interpersonal comparisons, whether through the "hypothetical compensation test" of Kaldor and Hicks or through the focus on "primary goods" in Rawls or "capabilities" in Sen.⁵⁶

That latter strategy of making interpersonal comparisons via something akin to welfare interests is particularly apt.⁵⁷ The objection to

⁵² The following paragraphs draw loosely upon my previous discussion of this topic (Goodin 1982b, pp. 16-8; 1991c, pp. 245-6).

⁵³ That is just the most dramatic case, though. The same problem arises if different courses of action benefit all, but differentially so. Then in choosing which is the utility-maximizing course of action we would once again need to compare those differential benefits across people.

⁵⁴ The old-fashioned hedonistic phrasing here is deliberate, for this is a criticism that dates primarily to the days when utilitarianism was essentially a hedonistic matter of maximizing pleasure and minimizing pain.

⁵⁵ Jevons 1911, p. 14. Robbins 1932, pp. 122-5; 1938.

⁵⁶ Hicks 1939. Kaldor 1939; 1946-7. Rawls 1971, pp. 90-5. Sen 1982a, pp. 203-4; 1985a.

⁵⁷ Indeed, it represents a return to the older "material welfare" school of Marshall and Pigou, wherein "the comparison of needs, not the comparison of subjective desires, was what they usually meant by comparing utilities of different people" and their focus was accordingly on observable indicators of "industrial efficiency," such as inadequacy of people's diets and measures of mortality and morbidity (Cooter and Rappoport 1984, p. 516).

interpersonal comparisons was only ever particularly compelling against now abandoned forms of crude hedonistic utilitarianism, anyway. To compare pleasures and pains across individuals, we may well need to get inside one another's heads. But that is unnecessary, once we have shifted to talking of what counterfactually *would* give them pleasure or pain. To answer questions such as these, we necessarily abstract from what actually does give them pleasure to what would (more food, healthier lives and so on).⁴⁸ Having abstracted in that way from actual preferences to hypothetical preferences – from desires to welfare interests – there ceases to be any barrier of a strictly meta-physical sort to comparing those abstractions across individuals. We are working well outside their heads already.⁴⁹

Practical people engaged in the real business of the world have, in any case, always found it difficult to take particularly seriously such worries about interpersonal utility comparisons. Typical of their reaction is that of Pigou, who scoffs, "Nobody can prove that anybody besides himself exists, but, nevertheless, everybody is quite sure of it."⁵⁰

There are various ways of making these worries look more than a little silly. Perhaps most tellingly, public debate, deliberation and even conversation itself would on the selfsame grounds prove quite impossible. If the problem with interpersonal comparisons of utility is that we cannot look into other people's minds to see how they feel, then there is a perfectly parallel problem with looking into their minds to see what they mean when uttering a statement. That does not stop us from talking – or even, after a fashion, conversing. We simply employ a "principle of charity" assuming that they, like us, try to talk sense and we ascribe meaning to their utterances accordingly.⁵¹

In trying to make sense of interpersonal utility comparisons, we need do no more than that, either. We need simply assume that others are much like ourselves, and act accordingly. Interpersonal interac-

⁴⁸ Some would regard this as doubly hard: not just a matter of getting inside people's heads to surmise their pleasures and pains, but doing so counterfactually (what would give them pleasure or pain). But the virtue of the counterfactualization here is to take us away from focusing upon people's heads and toward properties of goods. Surely we can abstract in this way, saying something in general about the qualities in commodities that give rise to satisfaction for any particular person. As Lancaster (1966; 1971) says, it is surely no mystery that people tend to regard margarine as a good substitute for butter but not for a motorcar.

⁴⁹ Goodin 1991c, pp. 245–6.

⁵⁰ Pigou 1951, p. 292. Cf. Little 1957, chap. 4; Barry 1965, pp. 44–7.

⁵¹ Davidson 1986.

tions could not proceed otherwise. Neither, more especially, could politics.⁶²

Such procedures will, perhaps inevitably, give rise to comparisons of utility that are only rather rough and ready. But perhaps it would amount to spurious precision to strive for more, given the various other uncertainties and ambiguities that characterize the context in which public policies are made. It would be folly to strive for accuracy to the fourth decimal place in one element of our utility sums, when other elements of it are reliable only to the first. And it is all too often the case that ballpark figures and order-of-magnitude estimates are often the best we can do in public policy-making.

That concession, if true, means that there will inevitably be some (perhaps substantial) imprecision in our public utility calculations. It follows that utilitarian policy recommendations will therefore still be somewhat indeterminate. That is not necessarily a telling criticism of utilitarianism, though. It merely amounts to saying that utilitarianism leaves some room for public debate and deliberation – in other words, for politics as it is ordinarily conceived. We would, I think, worry about any political theory that did not do that.⁶³

Furthermore, even where utilitarianism proves indeterminate, it sets the terms of that public debate. It tells us what sorts of considerations ought to weigh with us, often while allowing that how heavily each of them actually weighs is legitimately open to dispute. Even where utilitarianism is indeterminate, it is not silent. To fill in those lacunae, we do not need to turn to some other principles altogether. Rather, in such cases utilitarianism speaks with many voices, and political argument in such a setting can (utilitarians would say “should”) consist simply in a debate among them.

Once we have at least rough-and-ready interpersonal comparisons on the table, we can resume discussing the utilitarian merits of alternative distributions of the national dividend. Of course, which distribution utilitarianism recommends naturally depends upon certain crucial empirical facts of the matter. That in itself has long been taken to be a criticism of utilitarianism. If the facts of the matter turn out one particular way, utilitarianism might find itself recommending distributional policies that are intuitively outrageous.

⁶² As Waldron (personal communication 1993) points out at this point, politics is and must necessarily be ultimately based on that which we all have in common.

⁶³ That is a more straightforward way of saying what is wrong with Sidgwick's (1874) 1907, bk. 4, chap. 5, sec. 3) behind-the-back version of “government house utilitarianism.” The problem is not so much that it violates some high-minded Kantian publicity principle as it is that it is a political theory that curiously cuts out politics altogether (cf. B. Williams 1973a, pp. 138–40; Sen and Williams 1982, p. 16).

Utilitarianism might recommend feeding of Christians to lions, if it so happens that the utilities of the spectators enjoying the show (plus that of the lions enjoying the meal) exceed the disutilities of the Christians being sacrificed. Or utilitarianism might recommend dissecting one person and distributing her body parts to various others in need of transplants, if it so happens that the utilities of the recipients exceed the disutility of the "donor."⁴⁴ Or utilitarianism might recommend the hanging of an innocent person to assuage an angry mob, if it so happens that the utilities of those spared the mob's wrath exceed the disutility of the hanging victim.⁴⁵ Or utilitarianism might recommend giving all resources to a handful of people, if it so happens that those people are "super efficient pleasure machines" capable of translating resources into satisfaction at a fantastic rate; or it might recommend giving no resources to the handicapped, if it so happens that those people are particularly inept at translating resources into satisfaction.⁴⁶

There is no denying that utilitarian prescriptions might turn out that way, in any particular instance. There is no telling how the numbers will come up in each and every case. But, again, advocating utilitarianism as a public philosophy spares us the burdens associated with maximizing at the margins in each and every case. It involves instead adopting institutions and practices and policies, on a utilitarian basis; and those must, by their nature, be publicly accessible and relatively long lasting. That in turn means that in choosing institutions and practices and policies we cannot maximize at the margins, adapting our choices to peculiarities of utility mixes in particular cases. We must instead adapt our choices to standard situations recurring across protracted periods, and do so in full knowledge that the nature of our choices will sooner or later become common knowledge.

That fact goes some way toward ensuring that utilitarianism, practiced as a public philosophy, will have few of the grievous distributional consequences commonly supposed. Many of the cases involving sacrificing the interests of the few to the many (or of the many to the few) generate the purported utilitarian payoffs only if it never becomes public knowledge what we are doing. Once it becomes public knowledge that, as a matter of policy, we are willing to hang innocent people to assuage a baying mob or to carve up one person to generate spare parts for others, then everyone starts worrying: Who will be next? The anxieties associated with such thoughts systematically occurring across the whole population will more than suffice to cancel

⁴⁴ J. Harris 1975.

⁴⁵ Mabbott 1939. Lukes 1993, p. 428.

⁴⁶ Friedman 1947. R. Dworkin 1981.

the utility advantages of carving up one person or throwing one prisoner to the mob on any given occasion.

Utilitarianism, employed as a public philosophy, must by its nature adopt institutions and practices and policies suited to recurring situations and standard individuals. There may be a very few people who are vastly better and a very few who are vastly worse than others at translating resources into utility. But if one assumes, with Bentham and his followers, that most people are pretty much alike in this respect – and, further, that most goods display “diminishing marginal utility” (so the more you have of something, the less utility you derive from the next increment of it you are given) – then the utility-maximizing distribution of resources is inevitably going to be a broadly egalitarian one.

Anti-utilitarians complain loudly and often that utilitarianism disregards the morally crucial fact of the “separateness” of persons.⁷ That complaint, fair enough in a way, is however untrue in two crucial respects. First, utilitarians regard each person as a distinct locus of value. In generating the utilities that end up being aggregated, “each counts for one and no one counts for more than one,” in Bentham’s famous phrase. Of course, in the process of aggregating, the boundaries between you and me, your utilities and mine, get lost. But, second, empirical assumptions of broad similarity among people and generally diminishing marginal utility across all resources lead utilitarians to embrace policies and practices and institutions that are broadly egalitarian in form. That ensures that there will be a strong utilitarian presumption against exploiting some people for the benefit of others.

VI

Replying to James Mill’s “*Essay on Government*,” Macaulay takes this swipe at “*Benthamites*” more generally:

[H]aving read little or nothing, [they] are delighted to be rescued from the sense of their own inferiority by some teacher, who assures them that the studies which they have neglected are of no value, puts five or six phrases into their mouths, lends them an odd number of the Westminster Review, and in a month transforms them into philosophers. . . . [T]hese are smatterers, whose attainments just suffice to elevate them

⁷ This theme – recurring throughout Rawls 1971, B. Williams 1973a, Nozick 1974 and R. Dworkin 1977 – is most effectively summarized in Hart [1979] 1983.

from the insignificance of dunces to the dignity of bores, and to spread dismay among their pious aunts and grandmothers.⁶⁶

Intemperate though his language may be, Macaulay clearly has a point. One of the great advantages of utilitarianism has always been that it promises to yield determinate, no-nonsense advice on practical matters of what one should do. One of its great disadvantages has always been that it has a tendency to do so (at least in the hands of its most brusque, no-nonsense practitioners) in a singularly formulistic way. List the alternatives, list the consequences, attach utility numbers to each and crank the adding machine's handle. Nothing could be easier. But, critics say (with some considerable justification), nothing quite so easy could possibly be right.⁶⁷

There is no denying that many of the applications of utilitarianism to problems of public policy are just as rote as that. In a way, though, it is a virtue of utilitarianism that it is an ethic which admits of rote learning of that sort. Better that an ethic be applied by rote than not at all, if (or where) those are the only options - as often they are, given the limits to policy-makers' time, attention and talents.

In any case, utilitarianism of the most formulistic sort is sometimes transparently the right way to approach a policy problem. Suppose we are trying to assess the economic effects of income transfer programs, for example. Then balance-sheet thinking is precisely what we need. The traditional complaint against generous income support programs is that if people can get something for nothing then they will not bother working for a living in the present or saving for the future. But the magnitudes here clearly matter. American evidence suggests, for example, that in exchange for a 4.8 percent reduction of labor supply (and a reduction in private savings of between 0 and 20 percent) we get a 75 percent reduction in poverty and a 19 percent increase in equality (measured by the Gini coefficient).⁶⁸ Whether we think on balance the gains are worth the costs is an open question. That depends on the relative weights we attach to each of those factors. But whichever way we go on that concrete case, listing the various effects and weighing them against one another surely is the right way to go about making that an economic assessment of that sort.

Transparently right though such formulistic approaches to policy puzzles sometimes are, however, it would be wrong to judge utilitar-

⁶⁶ Macaulay [1829] 1978, pp. 99-100.

⁶⁷ Hampshire's ([1972] 1978, p. 1) comment about utilitarianism having "now become an obstruction" to enlightened policy seems primarily a reference to the way in which the same utilitarian cost-benefit logic that served the Kennedy Administration so well in reforming the Defense Department (Hitch and McKean 1960) led to so disastrous consequences in Vietnam (Halberstam 1969).

⁶⁸ Danziger, Haveman and Plotnick 1981, p. 1019.

ianism wholly in light of them. In coming to an overall assessment of utilitarianism as a public philosophy, it would be wrong to fixate exclusively upon the most formalistic derivations of its least imaginative practitioners. We should attend at least as much to the more creative uses that can be made of the tools with which utilitarianism provides us, to possibilities that arise from working "in the shadows of utilitarianism," in Hart's phrase.⁷ In the examples that follow, I attempt time and again to show how utilitarianism's central concepts might, given certain features of the problem at hand, yield determinate policy advice – without resorting to simpleminded, and often simply impossible, cranking through the formula to reach a direct determination of what is likely to maximize sum-total utility.

Thus, in the example of Chapter 17, it is possible to say in the spirit of utilitarianism that unilateral nuclear disarmament would have been a good policy in an essentially bipolar world – not because that would maximize utility (absent probability numbers, that is a sum that cannot be done), but rather because it would make a modal change in the possibility of truly awful outcomes. It is possible to say, in the example of Chapter 14, that an unconditional income guarantee (negative income tax, basic income, call it what you will) is a good thing – not because that would necessarily maximize overall social utility in the presently prevailing circumstances, but rather because that policy would be minimally sensitive to shifts in prevailing social circumstances which always change far more rapidly than social policy. For that reason, unconditional income guarantees would be more likely to maximize utility across that wide range of changing circumstances. Or, in the examples running through Chapters 10 to 13, policies to buffer people against radical changes to the course of their lives would be a good thing – not because those are the most satisfying lives that people might live, but rather because the chopping and changing required to get to something else would be profoundly disruptive of what people find ultimately satisfying in their lives.

Of course, the bottom line in all those cases is that the policies are justified because ultimately they are utility-maximizing in some sense or another. Mine would hardly be a utilitarian theory at all, were it otherwise. Invariably, though, those are judgments made employing the apparatus of utilitarianism but without having recourse to fine-grained calculations of sums. The considerations that are deemed decisive there for policy questions are indisputably utilitarian-style considerations, bearing directly upon the preference satisfaction

⁷ Hart [1979] 1983, p. 222. He may be stretching the point too far to say that even Nozick and Dworkin are working "in the shadow of utilitarianism" searching for grounding for their theories of rights – perhaps its shadow does not extend quite that far.

(somehow conceived) of people (somehow specified). The point just is that those considerations can indeed prove determinative as regards utilitarians' policy recommendations, well ahead of doing a full-dress utility count.

My concern in this book, true to the thrust of this introduction, is with utilitarianism as a public philosophy. My main concern is with the ways in which utilitarianism can be a good guide to public policies, without necessarily being a good guide to private conduct. Nonetheless, in adducing many of its most important implications for public policy it is important to see at least in broad outline how it would set about shaping private conduct.

Utilitarians, and consequentialists more generally, are outcome-oriented. In sharp contrast to Ten Commandment-style deontological approaches, which specify certain actions to be done as a matter of duty, utilitarian theories assign people responsibility for producing certain results, leaving the individuals concerned broad discretion in how to achieve those results. The same basic difference in the two theories' approaches to assigning moral jobs reappears across all levels of moral agency, from private individuals to collective (especially state) actors. The distinctively utilitarian approach, thus conceived, to international protection of the ozone layer is to assign states responsibilities for producing certain effects, leaving them broad discretion in how they accomplish it (Chapter 18). The distinctively utilitarian approach, thus conceived, to the ethical defense of nationalism is couched in terms of delimiting state boundaries in such a way as to assign particular responsibility for every particular person to some particular organization (Chapter 16). And, at a more domestic level of analysis, the distinctively utilitarian approach to the allocation of legal liabilities is to assign them to whomsoever can best discharge them (Chapters 5 through 7).

The great advantage of utilitarianism as a guide to public conduct is that it avoids gratuitous sacrifices, it ensures as best we are able to ensure in the uncertain world of public policy-making that policies are sensitive to people's interests or desires or preferences. The great failing of more deontological theories, applied to those realms, is that they fixate upon duties done for the sake of duty rather than for the sake of any good that is done by doing one's duty. Perhaps it is permissible (perhaps it is even proper) for private individuals in the course of their personal affairs to fetishize duties done for their own sake. It would be a mistake for public officials to do likewise, not least because it is impossible. The fixation on motives makes absolutely no sense in the public realm, and might make precious little sense in the private one even, as Chapter 3 shows.

The reason public action is required at all arises from the inability

of uncoordinated individual action to achieve certain morally desirable ends. Individuals are rightly excused from pursuing those ends. The inability is real; the excuses, perfectly valid. But libertarians are right in their diagnosis, wrong in their prescription. That is the message of Chapter 2. The same thing that makes those excuses valid at the individual level – the same thing that relieves individuals of responsibility – makes it morally incumbent upon individuals to organize themselves into collective units that are capable of acting where they as isolated individuals are not.

When they organize themselves into these collective units, those collective deliberations inevitably take place under very different circumstances, and their conclusions inevitably take very different forms. Individuals are morally required to operate in that collective manner, in certain crucial respects. But they are practically circumscribed in how they can operate, in their collective mode. And those special constraints characterizing the public sphere of decision-making give rise to the special circumstances that make utilitarianism peculiarly apt for public policy-making, in ways set out more fully in Chapter 4. Government house utilitarianism thus understood is, I would argue, a uniquely defensible public philosophy.⁷⁴

⁷⁴ I am grateful for comments on earlier drafts of this chapter from Peter McCarthy, Doug MacLean and Jeremy Waldron.

The state as a moral agent

Positive arguments for state intervention of a broadly utilitarian sort face essentially two different foes within contemporary moral and political philosophy. On the one hand, there are libertarians opposed to state intervention of any sort as a matter of principle. To defeat their objections, the utilitarian needs to show there is a legitimate role for state action at all. On the other hand, the utilitarian fixation with securing good outcomes through social action is opposed by deontologists, who would judge all actions in terms of the inputs – specifically, in terms of the quality of the motives and intentions from which they proceed. Defeating these two sorts of opponents is the task of this and the next chapter, respectively, thus paving the way for the introduction (in Chapter 4) of my preferred form of “government house” utilitarianism.

The essentially antilibertarian argument of this chapter starts from the observation that individuals find all sorts of excuses for not doing the right thing. What follows is an analysis of how two such excuses cut across each other in some rather surprising ways. The failure of isolated individuals to do the right thing is excused, but only at the price of permitting (or perhaps even requiring) coordinated collective action wherein individuals may be rightly required to play their respective parts.

The analytic key to this argument is simply the proposition that the state is a moral agent, too. It has responsibilities of its own (or, if you prefer, we have responsibilities through it) even where – indeed, especially where – we as individuals are excused from any responsibility for undertaking isolated, independent action. Hence the basic finding of this chapter – arguments for letting individual moral agents off the hook have the effect of putting collective moral agents such as the state on it, in their stead.

I. TWO EXCUSES

How exactly we are to determine what is the "right thing" to do is problematic. Happily, that whole set of issues can be safely left to one side for the purposes of this chapter. My story here begins after that one is finished. Even after it has been agreed what the right thing to do would be – however that has been decided – there is still a variety of excuses that can be given for not doing what it is agreed that, in some sense, ideally we ought to do.

Two such excuses concern me here. One can be called the "libertarian" excuse it is right that something be done, but wrong to force people to do it. The other can be called the "no individual responsibility" (or, more colloquially, the "not my job") excuse. It is right that something be done, but it is not my job to do it; I have no individual responsibility in the matter.

Each of these excuses is plausible in its way. Take the libertarian excuse, for example. There genuinely are many things that we think it right that people should do, but wrong that they should be forced to do. One reason has to do with the crucial role of people's intentions among the right-making characteristics of the act. Sometimes it is thought not to be enough that the right thing be done; it must also be done for the right reasons, if it is to count as morally worthy. Locke gives one indisputable example in his *Letter Concerning Toleration*, when he points out that pious acts performed by non-believers merely to avoid social sanctions will not suffice to procure their salvation.¹ Kantians would place most acts in the same category. Where the right act must be performed for the right reason, if it is to be right at all, forcing people to perform it would be futile. If their reason for performing the act were merely to avoid social sanction, then they would be acting from the wrong reason in a way that undermines the rightness of the act itself. It is strictly impossible, in such circumstances, to force people to act rightly.

Another argument tending toward the same practical conclusion has to do with the value of liberty. In the preceding argument, right acts lose their rightness when performed merely to avoid social sanction. In this next argument, right acts retain their rightness even when forced; but force constitutes a moral cost all its own, to be set off against any moral gains achieved by securing superior performances. This story can be fleshed out in various ways. We might say that liberty is a moral end in itself, and that its being infringed is in itself a moral cost. Or (or "and": these are nowise incompatible options) we

¹ Locke [1689] 1946.

might take a "developmental" line, emphasizing the importance of letting people learn from their own mistakes, so that they make fewer of them in the future. Here, liberty would not be an end in itself, but rather a means of producing better outcomes at less cost in future.

The "not my job" excuse, similarly, often has a good deal of surface plausibility. There may be something paradoxical about admitting that it would be good for something to be done and yet, at one and the same time, denying that it would be good for you to do it. But the paradox is more apparent than real. In truth, there are various plausible reasons for admitting that it would be good for something to be done (by someone) and yet denying that it should be morally mandatory for you, in particular, to do it.

One reason might be that the good to be done is something that you yourself simply cannot do. Sometimes the impossibility is strictly logical. It may be good that parents care for their own children; but, not being one of this child's parents, that particular good is not one that can be attained by your caring for the child. Other times, the impossibility might be merely contingent. As it happens, you are incapable (or much less capable than someone else) of doing what is required to promote the particular good in view. Yet other times, the impossibility might be of a psychological sort. Given your own deep commitments (to people, projects, and so on), and given your own psychological makeup, you simply cannot bring yourself to do what might be required of you to promote the good wherever and whenever you could. There are some things that you might do that truly are above and beyond the call of moral duty; some of them are such that it is all right for you to omit performing them.²

This last proposition, in particular, relies on something like the "bottomless pit" objection. At root, the protest there is that, if you were morally required to do all things of this sort that would be morally good to be done, then too much of a sacrifice would be demanded of you. This might be due either to the depth or to the scope of those demands. Doing the right thing on any one occasion might be just too costly reasonably to demand it of a person; or, while doing the right thing on any one occasion would not be so very costly, there might be just too many occasions that are identical in all morally relevant respects for us reasonably to expect you to do the right thing in all of them.³

That latter thought is much of what motivates the standard distinction between "perfect" and "imperfect" duties. The former are strong demands – obligations to be discharged every time the situa-

² Urmson 1958. Heyd 1982. Pettit and Goodin 1986, pp. 651–9.

³ Fishkin 1982.

tion arises. The latter are weak demands, in comparison. You must discharge such obligations on at least some occasions, but you need not discharge them on all occasions when they arise. Contractual obligations are a standard example of the former, "duties of charity" of the latter. This distinction between perfect and imperfect duties has implications for the enforceability of such duties, in turn. Since we ought always discharge each perfect duty, it is morally proper to enforce such duties whenever they arise, in a way that it would not be appropriate to enforce each instance of an imperfect duty which we need not discharge every time it arises.⁴

II. EXCULPATING INDIVIDUALS, INCULPATING STATES

Those are very standard perspectives on a pair of very standard excuses. My aim here is neither to challenge nor to elaborate on them. It is instead to show how those two excuses can cut across one another. The first step in this argument lies in showing how the second ("no individual responsibility") excuse, while exculpating individuals, inculcates collectivities. The next step, in Section III, is to show how that in turn undermines the "libertarian excuse," justifying collective enforcement of moral demands upon individuals.

So let us consider, then, the excuse that something is "not my job" or "not my responsibility." Certainly there are good reasons (of a pragmatic, and perhaps even deeper, sort) for a moral division of labor. There are good reasons for assigning particular, morally important tasks to some particular moral agents as their exclusive realm of responsibility. Once that assignment has been made, there are good reasons (pragmatic, certainly, and perhaps even deeper) for not trespassing upon someone else's moral preserve. If that is what we mean by saying, "It's not my job" or "It's not my responsibility," then the force of that proposition is plain for all to see.

The form of the phrase "It's not my job" may well serve to suggest that that is what the excuse standardly means. Saying, emphatically, "It's not *my* job," might seem to carry with it the unspoken message that it is someone else's. But, of course, there is no reason for supposing that that will always be true. If responsibility has been assigned to no one, then it is true of each in turn that it is "not my job."

⁴ While it may be inappropriate to force you to ϕ on each of the n occasions when it would be possible for you to ϕ , it would nonetheless be appropriate to require you to discharge your imperfect duty of ϕ ing on at least some of those n occasions. The barrier to enforcement of such a duty is presumably purely administrative. It is simply harder to enforce a rule that you should sometimes ϕ than it is to enforce one that you should always (or never) ϕ .

In that case, much of the former force of the excusing proposition is lost. If it is neither your job nor anyone else's, then the objection to doing something that it is "not your job" cannot be that you will be meddling in someone else's moral business. It cannot be that you will be undermining the moral division of labor. And, while it is still perfectly possible, it is much less likely under those circumstances that you will be cutting across someone else who is trying to do the job himself. Thus, it is much less clear why you should be excused from doing something, just because you have not been assigned responsibility for it, when no one else had been, either.

As Austin says, excuses characteristically get us out of the fire into the frying pan. Getting off the hook in one way puts you on the hook in another.¹ Or getting someone off the hook puts someone or something else there, in his place. That is as much the case as ever with the "no individual responsibility" excuse. It gets individuals off the hook, but succeeds only in putting the group there in their place.

What is no one's responsibility is everyone's. If it is right that something be done, and no one in particular has been assigned responsibility for doing it, then we are all responsible for seeing to it that it be done.

To say that all are responsible is not necessarily to say that each is responsible, though. Still less is it to say that each is necessarily responsible for attempting to do whatever must be done himself. The act in question may be such that it is either unnecessary or even counterproductive for everyone to perform it. It may be the case that, once any one of them performed it, none of the others would need to; or it may be the case that, once any of them commenced performing it, others' attempting simultaneously to do so would constitute counterproductive interference. Or the "bottomless pit" problem, canvassed earlier, might arise. Asking everyone to perform all morally desirable actions might impose an unreasonable sacrifice on each of them. For such reasons, we typically – and rightly – suppose that, when responsibilities have not been allocated to anyone in particular within a group, the most that can be said is that each of them has an imperfect duty to perform at least some (but not necessarily all) of the acts that we might ideally wish be performed.

The same general principle gives rise to much stronger implications at the level of the group as a whole, however. When no one in particular bears responsibility for performing some morally desirable ac-

¹ Austin 1956–7, p. 3. Consider the excuse of impossibility by reason of avoidable ignorance or incapacity, for example. Being excused from doing the right thing now comes at the cost of being blamed for not having done something in the past that would have made you able to do the right thing later (Goodin 1982b, chap. 7; Zimmerman 1987).

tions, everyone collectively has a strong, perfect duty to see to it that those things are done, within the limits of the capacities of the group as a whole to do so without undue sacrifice.⁶

The argument for strong collective responsibility in such cases proceeds by two steps. First, notice that the problems posed here for individual action in the circumstances envisaged are all, in essence, coordination problems.⁷ Where two or more people try to do the same good deed, their efforts might prove counterproductive; or one's good deed might render the other's superfluous. Either way, there is a need for coordination. Similarly, the danger of bottomless pits arises principally because, in the absence of coordinating mechanisms, the conscientious rightly dread that the unconscientious will shirk their duties, leaving the former with much more than their share of the moral chores to do. Running through all the rationales for the "not my job" excuse, then, is this one common feature – they all point to coordination problems, of one sort or another.

The second step in this argument is to show that the solution to such coordination problems is, of necessity, a responsibility peculiar to the group as a whole. To some extent, this follows from the very nature of coordination. By its very nature, coordination is not something that can be performed by one actor in isolation. By its nature, "coordination" refers to the relations between things: between muscles, in a graceful dancer's body; between departments, in a smoothly functioning state; or, in the case of social coordination, between your actions and others'. Coordination simply cannot be an attribute of your action alone, or of others' actions alone. Rather, it can only be an attribute of yours *together with theirs*. And that, in turn, is the essence of a "group action."⁸

All that that argument strictly shows, of course, is that it is only in groups that behavior can be socially coordinated. It does not yet establish that it is the job of groups to coordinate behavior. Coordination, by its nature, is a collective enterprise. But it remains possible for the coordination to be accomplished by individuals' action *in*

⁶ Goodin 1985c, pp. 134–44; for applications, see Wasserstrom 1975, p. 9; 1983, p. 30. Of course, if moral duties far outstrip the resources of the group as a whole, then the whole group should be excused from them. But often pits that seem bottomless from the perspective of the individual are perfectly manageable from a collective perspective. If we can merely ensure that everyone contributes his fair share, we can manage the burden perfectly happily. The only reason the pit looked bottomless from the individual's perspective was that he was contemplating having (at least potentially) to do his bit and everyone else's as well. In this chapter, I concentrate on burdens of this latter sort – i.e., where the "bottomless pit" argument excuses individuals but not groups as a whole.

⁷ Lewis 1969. Goodin 1976, chaps. 4 and 5.

⁸ F. Jackson 1988.

groups. The collective nature of the enterprise is respected by stipulating "in" groups. But that does not imply that the only way they can coordinate behavior is through some heavy-handed, formal, collective action "by" groups.

Casual reflection on the coordination process clearly reveals that neither formal organization nor collective participation of the group as a whole is always strictly necessary for coordination to succeed. Quite the contrary. Coordination is sometimes best achieved by delegating responsibility for coordination to some particular member or members of the group. Other times it is best accomplished by assigning the task to someone outside the group altogether. Still other times, coordination might best be left to emerge naturally, through group interactions or through a shared perception of certain "obvious" points as the foci for concerted action. Any of these solutions might, in principle, produce the desired coordination."

Having said all that, however, there remain certain roles which groups necessarily have to play in coordinating behavior. First, where coordination does not emerge naturally, coordination schemes can function as coordination schemes at all only if they are embraced by the group whose behavior is to be coordinated by them. This, in turn, means that someone must intentionally have engineered the coordination scheme, and everyone must act intentionally in compliance with it.¹⁰ That is to say, coordination requires everyone to "track" everyone else's behavior. When one person's behavior changes, everyone else must take note of the fact and be prepared to make the necessary changes in their own behavior in response." Where this does not happen "naturally" (for instance, through the market), it can only happen intentionally - either directly (through the agent's own intentions) or at one remove (through the agent's intentional response to the system engineer's intentional designs).

Second, even where there is no need to organize a coordination scheme formally, the group as a whole still has a residual supervisory function. This entails, in the first instance, a responsibility to undertake regular monitoring. It entails, in the second place, a responsibility to be prepared to organize a more formal coordination scheme should less formal ones fail to perform satisfactorily.

Thus, groups must be at least *ultimately* responsible for coordination. The reason is the same as the reason why I must be responsible for my own sins, or why the sole onlooker must be responsible for

⁹ Goodin 1976, chap. 5.

¹⁰ People might just intend to avoid sanctions designed to enforce compliance, of course. So long as the scheme has been intentionally engineered in such a way as to guarantee that that is extensionally equivalent, that is enough.

¹¹ Nozick 1981, pp. 317-26.

rescuing the drowning swimmer – no one else can, or will. Coordination is, by its nature, *our* collective enterprise. No other agent, individual or group, can do it for us. If the behavior is ours, then any coordination it manifests must necessarily be ours – at least in the sense of being an attribute of our, rather than anyone else's behavior.

If it is good that we should coordinate our behavior in these ways, as *ex hypothesi* it is in the cases here in view, then we must ultimately be prepared to do so through our own collective efforts. If individuals are rightly to be excused from achieving the good through their own isolated actions, pleading "It's not my job," then the collectivity must be empowered and enjoined to do whatever is necessary to eliminate those barriers that block morally efficacious individual behavior. The collectivity must be empowered to *make it* someone's job, if anyone is to be allowed to plead, "It's not my job."

What it might mean to hold responsible a group as a whole, where the group is in no way formally organized, is perhaps unclear. Responsibility implies agency, and agency implies some capacity for intentional action. Unorganized groups lack that. Random collections of people (for instance, that group of people occupying the third carriage of the 17:27 train from London to Clacton) are incapable of forming any "collective intention" until they first form themselves into a properly constituted "collectivity"; so, too, are members of an unruly mob in the street. People might be held individually responsible for joining in the mob's rioting; or they might be held individually responsible for not doing whatever was required to constitute themselves into a proper collectivity, capable of rescuing others who are trapped in the carriage when the train crashes off the rails.¹⁴ But the group, as such, cannot bear any responsibilities until it is properly constituted.

Where there is some collective agency in existence, though, there is no problem in ascribing group responsibilities of this sort directly to it. The state is preeminent among such organized collectivities. Our paradigm of moral agency is essentially individualistic, to be sure. The natural person is our model. Only those things that are sufficiently like natural individuals – only those things that are possessed of clear values, goals and ends, and capable of deliberation upon and intentional implementation of action plans in pursuit of them – can count as agents at all, for moral purposes. It is only to them that moral injunctions can be addressed. The limits of their capacity for effective action mark the limits of our moralizing.

But artificially created agencies are agents, too. Most especially, the state is a moral agent, in all the respects that morally matter. It, like

¹⁴ Held 1970; 1972. French 1979; 1984. Goodin 1985c, pp. 134–44. Pettit and Goodin 1986, pp. 673–6.

the natural individual, is capable of embodying values, goals and ends; it, too, is capable (through its legislative and executive organs) of deliberative action in pursuit of them. The state is possessed of an internal decision mechanism (a constitution, and the processes that it prescribes) that mimics perfectly, for these purposes, that which is taken as the defining feature of moral agency in the natural individual. Without such mechanisms, the state would not be a state at all. It would lack the minimal organizational content required for that description to fit. With such mechanisms, the state is indisputably a moral agent, much like any other.¹¹

None of this is necessarily to suggest that collectivities – even organized ones – have responsibilities which are not reducible to the responsibilities of individuals comprising them. It can, of course, be argued that they do. Indeed, that interpretation seems to follow naturally from ascribing some independent moral agency to the organization as such.¹² But it is also at least arguable that all the responsibilities of the organization are, in the final analysis, responsibilities of individuals – both individuals in positions of responsibility within the organization, and individuals on whose behalf the organization acts.

Nothing in the present argument requires me to take sides in this dispute. My claim could, in line with the latter position, be read as simply maintaining that individuals are excused from doing the right thing under one set of circumstances but not another. Specifically, where doing the right thing requires coordinated social action, they are excused in the absence of mechanisms to provide that coordination. They are not excused in the presence of mechanisms to provide it. Nor are they excused from a duty to create and maintain those mechanisms to provide such coordination.

Neither is any of this to suggest that the state is the only collective agent capable of providing social coordination. Corporations, clubs, churches, and so on are all, in principle, capable of helping to coordinate individual behavior in much the same ways. All of them are, by reason of their internal decision apparatus, capable of being held responsible for doing so, in much the same way as the state. And so on.

Still, the state must be ultimately responsible, because the state is the preeminent organization among them in any given territory. Other organizations exist by leave of – and at least in one (legalistic) sense, only under a charter from – the state. Any sanctions that those other organizations want to impose upon their members in order to enforce

¹¹ French 1979; 1984. Held 1970; 1972.

¹² Held 1970; 1972. F. Jackson 1988.

a coordination scheme can ultimately only be imposed by leave of – and, often, only with the assistance of – the state, through its monopoly on legitimate violence. If coordination is required, other organizations might be responsible in the first instance for providing it. But, by its nature, the state must be the collective agency ultimately responsible.

Lest the basic structure of my argument get lost among the finer details, let me now summarize the argument so far. The key moves in the argument have been: first, to show that the validity of the “no individual responsibility” excuse follows from the impossibility of efficacious individual action, which in turn follows from the existence of a coordination problem of some sort or another; and, second, to show that, at least in the presence of formally organized collectivities, the existence of a coordination problem implies that the collectivity (and, among collectivities, ultimately the state) must bear ultimate responsibility for providing the coordination that is required in order for people to be able to do the right thing. In short, the same thing as makes it valid for individuals to offer the excuse of “no individual responsibility” implies that there must, in such situations, be a collective responsibility.

III. TAKE NO LIBERTIES WITH SHARED RESPONSIBILITIES

The next step in my argument builds on the results of the last. Once it is established that we have some shared, collective responsibilities to do something or to see to it that something is done, libertarian excuses can be shown to lose much of their force. The libertarian excuse is principally an excuse for not forcing people to discharge their *own* isolated, individual moral responsibilities. It does not work nearly as well (typically, it does not work at all) as an excuse for allowing individuals to refuse to play their part in discharging shared, collective responsibilities, and thereby to prevent us from discharging *our* moral responsibilities.

At root, the libertarian principle is one enjoining us to mind our own moral business. If a moral agent fails to discharge his own duties, that is his problem, not ours. It is wrong of him not to do what he should do, of course. But it is his wrong, not ours. It would be wrong of us to interfere in what is not, ultimately, any of our moral business by attempting to force him to do his duty.

Indeed, in the paradigmatic application of this libertarian principle it is doubly “his own business,” for the duty which we are excused from enforcing upon him is one of a peculiarly self-regarding kind. Not only would it do no good for a person to be forced to worship a

god in which he does not believe, but he alone would suffer from his delict.

Not all applications of the libertarian principle point to duties that are quite so narrowly self-regarding, of course. Other people often will suffer as a result of the agent's moral delicts, being harmed in one way or another. The point of the libertarian principle is not to deny that this ever happens; it is instead to emphasize that, when it does, the blame falls squarely upon the individual whose duty it was to do those things. It was his job to do them, not ours to make him do them. If he has wronged others through his delicts, the wrongs are wholly chargeable to his moral account. It is nothing to do with any of the rest of us. Any writs that are issued in consequence of the delict should be addressed to him, not us. In that sense, at least, even other-regarding acts are his business, not ours.

All that is meant merely by way of explicating, not of defending, the libertarian principle. Whether or not it is ultimately defensible, even as applied to isolated, individual moral responsibilities, I propose to leave here as an open question. My aim here is merely to show that, whatever you might think of that principle as applied to isolated, individual moral responsibilities, it is clearly inapplicable to shared, collective moral responsibilities.

Where shared, collective responsibilities are concerned, it is – by definition – everyone's business what everyone else does. And this tautology is far from an empty one. It is everyone's business, first and most simply, because it is a responsibility that everyone shares with everyone else. It is everyone's business, second and more importantly, because, for anyone else's contribution to be efficacious, each agent must usually play his part under the scheme that has been collectively instituted for discharging that shared responsibility.

When isolated, individual moral responsibilities are not discharged, various other people might be harmed. But at least no one else will (necessarily, or even usually) thereby be prevented from discharging his *own* moral responsibilities. When an individual fails to discharge responsibilities assigned to him pursuant to some scheme for discharging shared, collective responsibilities, this is not the case. The success of others' acts pursuant to such schemes will indeed typically be predicated upon the success of his own.

The need for coordination of this sort was what excused isolated individuals from acting on their own, and what in turn gave rise to the collective responsibilities in the first place. The failure of any one party to abide by the coordination scheme will typically undermine, to some greater or lesser extent, the success of the scheme as a whole, thereby preventing other moral agents from successfully discharging their assigned duties. It is for that reason that we may rightly force

people to do their duties pursuant to schemes for discharging shared, collective responsibilities – even if we may not so enforce isolated, individual responsibilities.

The rather more grand way of phrasing the point here might be couched in terms of undermining moral agency. Failure to discharge isolated, individual responsibilities may well result in other people's being harmed. That is wrong. But it is, at least in principle, a remediable wrong. People can, at least in principle, always be compensated for harms to their interests (or so the libertarian would claim, anyway). Failure to discharge shared, collective responsibilities has more grievous consequences, undermining in certain crucial respects other people's moral agency itself. For that, compensation is in principle impossible. There must be a moral agent to be compensated, and it is that very moral agency that is being undermined.¹⁵

A less grandiose, and more satisfactory, way of putting the point would be this. It may well be possible to compensate people, in *some* sense or another, for preventing them from doing their moral duty. Certainly there must be some sum of money large enough to make it up to them, in the sense that the people concerned would think themselves globally better off being very rich delinquents than they would have been as very poor saints. But global well-being is not the right standard here. Moral delinquents are *worse people*, not just *worse-off people*.¹⁶ Morally conscientious people keep separate accounting categories, entering moral credits and material payoffs in completely distinct columns. To say that they might be compensated, in the sense of being made globally better off by some transfer of money, is not to say that the surplus in the material-reward column can wipe out the deficit in the moral-virtue column. They may be better off with compensation, but they are differently off. Morality cannot be traded off for anything but morality. That is what explains the impossibility of compensating other people for the wrong done to them when they are prevented from discharging their collective moral responsibilities. That is what justifies us, *pace* libertarian principles, in forcing people to play their part in collective moral enterprises – so that others may play their part in them, too.

¹⁵ So too with murder and grievous bodily harm, perhaps. What justifies us in prohibiting those acts *ex ante*, rather than merely requiring compensation *ex post*, may well be that compensation there is in principle impossible. Clearly in the case of murder (and to some extent with gross bodily harm as well) the impossibility of compensation can be traced to the (total or partial) destruction of the moral agent who would have to have been compensated. On these issues, see Chapter 11 more generally.

¹⁶ True, when someone is prevented from doing his duty, the delict is not his fault. But someone who is relieved rather than resentful that others have spared him his duty would hardly count as morally conscientious.

All that is then left to underwrite the libertarian argument for not forcing people to do what they should do is the "value of liberty." In the circumstances here envisaged, though, the value of liberty weighs on both sides of the scale. We infringe a person's liberty by forcing him to do his moral duty, to be sure. But we also allow him to infringe others' liberty (i.e., their liberty to play an effective part in some collective scheme for discharging their shared moral obligations) if we allow him to default on his own obligations under that scheme. In such circumstances, we should not be surprised if the liberty of the many is characteristically taken to outweigh the liberty of the one.¹⁷

All of this is simply to say that, where there is a collective responsibility to coordinate individual behavior in pursuit of some morally important goal, it is legitimate for the collectivity to impose sanctions upon individuals in pursuit of that goal. Of course, it is perfectly true that not all coordination schemes require such enforcement. As I have acknowledged in the previous section, people are sometimes prepared to play their assigned roles without any external sanctions whatsoever. So my argument here is not that we should necessarily always enforce coordination schemes. It is, rather, that we should always be prepared to enforce them as necessary.

Neither is it necessarily the case, even where enforcement is required, that it is always essential to the success of a coordination scheme that everyone's compliance be secured. Often, we can afford a few defectors without any serious loss to the overall success of the project; indeed, insofar as possible, it is only sensible to engineer backups and redundant systems into coordination schemes so as to guarantee that this is the case. But, even where we could afford a few defectors, we nonetheless ought to impose the same sanctions against all of them alike. On this conclusion, considerations of fairness and prudence converge. If we were to set a differential tariff, depending on the actual harm that any particular defection caused, then everyone would be jockeying to be among the "affordable" defectors. There is, then, a grievous risk that too many would end up defecting, and that the success of the coordination scheme would be compromised in consequence.¹⁸

¹⁷ It is of course easy to justify infringing your liberty where you have explicitly consented to a coordination scheme with the enforcement provisions that are now being used against you. But even where you have not consented, infringing your liberty might still be justifiable. If you have a moral duty that can only be discharged through coordinated action, and that coordination can only be achieved through a scheme with enforcement provisions of that sort, then you have a duty to consent to such a scheme and its attendant enforcement. That duty to consent, rather than your actual consent, is what justifies levying sanctions.

¹⁸ Taylor and Ward 1982. Pettit 1986.

None of these arguments in defense of collective enforcement of coordination schemes to discharge collective responsibilities applies peculiarly to the state. All collectivities have a similar right to sanction their members for noncompliance with coordination schemes. States fine and imprison, seize and sequester; but so too do clubs fine their members, corporations demote or dismiss their officers, and so on. Still, all of that private sanctioning is conducted within the framework provided by state authority, and it is that which ultimately stands behind all sanctions levied by lesser organizations. So, in that sense, the state has to be the ultimate source of sanctions. In that sense, it is toward the state's sanctioning powers that these arguments in justification are primarily addressed.

Neither do any of the arguments offered so far specify the limits to the sanctions that may legitimately be imposed in pursuit of social coordination. It surely cannot be legitimate to levy any sanction, however large, in pursuit of any collective enterprise, however trivial. Some sorts of coordination clearly do matter more than others, however. And logically we should surely be able to levy larger sanctions in support of more important goals.¹⁹

Again, let me summarize the basic logic of my argument lest the detail obscure the basic structure. The argument of the previous section established that the same thing that makes the "no individual responsibility" excuse work to exculpate individuals from responsibility also works to inculcate collectivities, imposing upon them responsibilities to act so as to provide the needed coordination of individuals' behavior. The argument of this section has established that, whereas compelling people to do their individual moral duties might be impermissible, it is perfectly permissible to compel people to play their necessary parts in discharging collective responsibilities. That permissibility of compulsion arises from the fact that delinquents actually hinder others from discharging their own responsibilities under a coordination scheme. That is what makes it others' business what the delinquents do or fail to do; that is what justifies others in compelling the delinquent to play his necessary part in the coordination scheme. In short, the "no individual responsibility" excuse undercuts the libertarian plea of "no legitimate compulsion." You cannot have it both ways. Where one claim has force, the other is for the same reason deprived of its.

¹⁹ To make our sanctions socially efficacious, we must at least levy a sanction sufficient to extract from defectors the illicit gains they secure from defecting when others are cooperating. Where uncertainty surrounds detection and punishment, the sanction for those who actually are caught and punished needs to be proportionately higher in order to make the statistically expected payoff of defection the same. See more generally Goodin 1976, chap. 4.

IV. APPLICATIONS

For some indication of the practical import of this argument, consider two applications. One has to do with the state's responsibility for its citizens' physical security, and its right to employ coercive sanctions pursuant to the exercise of that responsibility. The other has to do with the state's responsibility for its citizens' economic security, and its right to employ coercive sanctions in pursuit of that.

This pair of applications is nowise unique. Various other examples could have been offered instead. But the pair I have chosen is particularly significant. Between them, "providing for the common defense" and "promoting the general welfare" represent two of the most central spheres of state activity. More to the strategic point of the present essay is the fact that one of these activities (promoting physical security) is broadly embraced by libertarians, whereas the other (promoting economic security) is stoutly resisted as a legitimate area of state responsibility. Showing that state responsibility – backed, indeed, by compulsion as necessary – can be justified in precisely the same manner in each of these cases will incidentally serve to undercut the libertarian case against compulsory state measures directed at promoting economic security.²⁰ That, too, is illustrative of the range and significance of the argument I have here been mounting.

Take first the case of physical security. Each individual is, of course, morally bound to refrain from attacking anyone else; and it is hard to see how the excuses here in view could ever release him from that duty. But each individual is also obliged to help protect others from attack. This, typically, is a duty that does call forth the sorts of excuses discussed here under the "no individual responsibility" heading. Sometimes the cost of protecting another's life or property would be the grave endangering of your own – especially if you are the only one to come to the other's aid. Sometimes it is clear that someone should help, but it is unclear who among several equally eligible candidates should render the assistance. Sometimes the desired protection can be accomplished only through a complex series of tasks that must be performed by different people, and your contribution would be useless (or worse) unless you could be sure that others would play their parts.²¹ And so on.

Because of the need for coordination in all these respects between people's efforts at protecting one another against physical assault, we typically excuse people from any individual responsibility in such

²⁰ This strategy is deployed with great success by Shue (1980, chaps. 1 and 2).

²¹ Cf. Held 1972, pp. 114–15.

matters. At most, "Good Samaritan" laws might impose a "duty of easy rescue" upon people, when it is clear what they could do to help and when rendering such assistance would be virtually costless for them.²²

But for the same reason that we ordinarily excuse individuals from responsibility for protecting one another's physical security, we typically (and rightly) deem it a collective responsibility. It is a duty of the group as a whole to create and sustain some scheme for coordinating its members' efforts at protecting each other's physical security; and, once some such collective scheme has been organized, the group has the right to use such force as necessary for compelling compliance with that coordination scheme, within the limits set by its nature as a coordination scheme. (That is merely to say that it must not exercise such force against anyone as to leave that person worse off than he would have been in a "state of nature," where there was no such scheme in operation at all and everyone was free to attack him.) As regards questions of physical security, all this is fairly well accepted, even by libertarians of Nozick's ilk.²³

The same sort of argument can also be made, however, as regards economic security. There, too, the "no individual responsibility" excuses are commonly invoked. Charity may be a moral duty, but it is an imperfect one. There are simply so many people in need that I would impoverish myself trying to cure poverty single-handedly. Besides, it is a complex issue, and there is no way for isolated individuals to be sure that their well-intentioned acts of private charity are not counterproductive.²⁴ And so on.

This, too, points to the need for coordination between people's duty-bound efforts at protecting one another's economic security. On account of that, we typically excuse people from individual responsibility in such matters. But, again for the selfsame reason, we ought to impose collective responsibility in these matters.

Here, again, if we take seriously the proposition that individuals have a moral duty (albeit only an imperfect one) to protect others' economic security, then it is their duty to create and sustain some scheme for coordinating aid-giving efforts within their group. Once some such collective scheme has been organized, the group has the right to use such force as necessary for compelling compliance with that coordination scheme, again within the limits set by its nature as a coordination scheme. Here that merely means that it must not ex-

²² Weinrib 1980. Kleinig 1976.

²³ Nozick 1974.

²⁴ McKinsey 1981.

ercise such force against anyone as to leave him less able to discharge his aid-giving duties than he would have been in the absence of such a coordination scheme in the first place.

This is a conclusion that libertarians would resist strongly, of course. They maintain that it is right to give, but wrong to be forced to give, charitable relief to those less fortunate.⁵⁵ But the argument here is perfectly parallel to that developed above for enforceable collective protection of physical security. In that form, it is an argument which they regularly welcome.

V. CONCLUSION

The upshot of this argument is thus powerfully antilibertarian. The conclusion is that the state has the duty to organize – and the power to enforce, as necessary – various sorts of coordination schemes to aid its citizens in discharging their individual (albeit imperfect) moral duties.

The larger target is moral shirkers more generally, though. People regularly try to find excuses for not doing the right thing. Some of the excuses they offer do indeed seem compelling. The argument of this chapter, however, is that, in their frenzy to excuse their delicts, these moral shirkers tend to trip over their own excuses.

Excuses rarely excuse completely. Some of them cut across one another in surprising ways. As I have shown here, we can sustain shirkers' claims that people have "no individual responsibility" in certain matters only by admitting that collectivities (such as the state) do have responsibility in those matters, and indeed that they have powers of compulsion as necessary to discharge those responsibilities. Dedicated moral shirkers will, of course, derive little comfort from such findings. But that is all to the good.⁵⁶

⁵⁵ Nozick 1974, pp. ix–x. Waldron 1986, p. 466.

⁵⁶ I am grateful to Alan Hamlin and Philip Pettit for comments on this chapter.

PART II

Morality, public and private

Do motives matter?

Utilitarian schemes for state intervention to promote the common good are opposed not only by libertarians opposed to state intervention as a matter of principle but also by deontologists opposed to the utilitarian fixation with good outcomes. What matters much more to them are individuals' motives and intentions. It is not enough, for them, that the right thing be done. They also insist that it be done, and be seen to be done, for the right reasons.

Thus, for example, deontological moralists and social critics under their sway are anxious to know whether we are sending food to starving Africans out of genuinely altruistic concern or merely to clear domestic commodity markets, for one particularly topical example.¹ Or, for another example, critics of the Brandt Commission's plea for increased foreign aid more generally say, in stinging rebuke: "Many of those who support the proposal . . . do so out of genuine humanitarian concern about . . . poverty. But it is doubtful whether this is the main concern of its authors, and it certainly is not their only concern. . . . They are, instead, primarily concerned with the preservation of the existing world economic order."²

What is common to all such cases is an attempt at motive differentiation. Any particular piece of behavior *might* have sprung from any of a number of different underlying motives; commentators (moralists, social critics) want to know which was the *real* motive. Here I shall show that this characteristic quest for motive differentiation is misguided. In most of the standard social situations, it makes no material difference to agents' actions whether they act from one sort of

¹ Saylor 1977, p. 202.

² Hayter 1981, p.9. W. Brandt (1980, p. 64) admits as much, saying, "It is a mark of the uneasy relations between North and South that even to speak of mutual interest can cause suspicion."

motive or another. And in such circumstances, pressing the motivational issue will usually lead only to mischief, of both a pragmatic and a moral sort.

1

The simple observation on which this argument builds is just this. Most people most of the time have very many different reasons for performing one and the same action. Bishop Tutu goes to Oslo to collect the Nobel Peace Prize. Why? Well, there are various reasons we might mention. It is a high honor and a lucrative purse. Those seem to be egoistic considerations, at least at first blush. But it is also a rare chance to speak out against the evils of apartheid from an enormously influential platform, perhaps thus hastening an end to the suffering of oppressed peoples across South Africa. That looks like a principled, altruistic concern on almost any reading. Which, you ask, was the bishop's *real* motive? What, I ask in replying on the good bishop's behalf, does it matter? Either way, he wanted to make sure to catch his plane to Oslo.

I take this to be an absolutely standard social situation. Putting the point in terms (which will serve merely as a useful shorthand) of choices being determined by reasons, my claim is simply that our choices are characteristically overdetermined. Since there are *more than enough* reasons for an agent to ϕ , there is no (direct, immediate, pragmatic) reason for the agent to inquire how many of those reasons would have been *just enough* to induce him to ϕ . The pragmatic question before him is merely whether or not to ϕ in the circumstances he is actually confronting. For him, and from that perspective, it is idle to speculate whether or not he would be similarly inclined to ϕ in some slightly altered counterfactual circumstances, where one or more of those reasons for action were removed.¹

What may tend to obscure this simple point is that there usually is a multitude of reasons for ϕ ing and a multitude of reasons for not- ϕ ing. An agent must choose either to ϕ or to not- ϕ . In so doing, he must decide which set of reasons weighs more heavily with him. That he must come down on one side rather than another of this question might, then, seem to suggest that he will have had to get clear about the relative strengths of all his different motives.

Nothing could be further from the truth, however. In deciding

¹ This is different from economists critical of revealed-preference methodology saying that a person's true motives might be empirically undecidable by external observers (Sen 1973) and from sociologists-cum-psychologists saying that a person's decisions are dictated by unconscious forces (Maclver 1940; Peters 1956).

whether or not to ϕ , an agent need only determine whether reasons R_1, \dots, R_m , taken together weigh more heavily with him than reasons R_m, \dots, R_n taken together. There is no need in the world for him to assign weights to each particular reason within the larger bundle of reasons, all of which argue in the same direction. Nor, having decided which set of reasons is on the whole more compelling, is there any need for the agent to determine *how much* more compelling it is. If there are *more than enough* reasons on the one side to outweigh those on the other, there is no need for the agent to inquire how many of those reasons would have been *just enough* to outweigh those on the other.⁴

II

To discover how common such a superfluity of reasons might be, notice that accusations of "rationalization" (which are very common indeed) presuppose just such a situation. When labeling someone's account of his actions a rationalization, we are implicitly conceding that those *would have been* good reasons for acting as he did; we are merely denying that those really were *his* reasons for acting as he did. We quite agree that behaving as he did would have been a good way of serving the goals he is now nominating; we merely deny that this post hoc reconstruction accurately reflects his thinking (and, specifically, his goals) at the time.⁵ His account would not be a rationalization – it would not "rationalize" his actions at all – were those not truly reasons for someone who did embrace the nominated goals to behave in the fashion the agent in fact did. Rationalizations misrepresent external reasons as internal ones. But the fact remains that there had to be multiple reasons, of one sort or the other, for the agent to do what he did in order for the allegation of rationalization to make sense at all. Hence, if we frequently have cause to fear rationalizations, then we must by the same token frequently have cause for supposing that there frequently exist multiple reasons for someone to perform one and the same action.

⁴ Economists face similar problems in estimating the "consumer's surplus" (Marshall 1920, bk. 3, chap. 6). It is enough that consumers know that they want to buy at the asking price; they have no reason to bother contemplating just how much more they would have been prepared to pay. Such considerations led Sidgwick ([1874] 1907, bk. 3, chap. 12, sec. 3) to despair of the "moral judgment of motives" altogether.

⁵ Notice that talk of motives typically does arise principally in after-the-fact justifications of one's actions to others (Peters 1956; Mills 1940).

There are several ways in which it might happen that we come to have such a superfluity of reasons for performing the same action. Here I shall identify three, two general in form and one peculiarly sociopolitical. All suggest this will be a common phenomenon.

Perhaps the most general ground is this. One and the same spatio-temporal sequence admits of many different descriptions. In the hackneyed example, someone raising his hand in some particular way might be waving, or saluting, or volunteering, or swatting a fly, or all of the above. Presumably different explanations will have to be given for – and hence different reasons will have to be appealed to, in order to justify – one and the same performance, depending upon whether we consider the action under one description rather than another. So, at this most general level, the multiplicity of possible descriptions guarantees a multiplicity of possible reasons for performing “one and the same action,” naively understood.⁶

Connected to that is a second general ground for supposing that there will ordinarily be a multiplicity of reasons for any given action. Before we can ask *why* a person did something, we must first ask *what* he did – what he saw himself as doing. Most people most of the time would describe any particular piece of behavior as falling under some more general rule, practice or “habit” which they have adopted for guiding a wide variety of similar activities.⁷ Very few people indeed decide each case completely on its own merits.

If people’s choices of actions are self-consciously rule-bound, then the question becomes one of what motivated those people to adopt those particular rules. The answer, naturally, varies from agent to agent and rule to rule. But what matters in the present context is just this – these more general rules will, ordinarily, offer more numerous and more diverse reasons for being followed. Hence, even if in any particular instance there is only one possible reason for an agent to perform any particular action, the agent may nonetheless see himself as acting on a more general rule which there are multiple reasons for respecting.

⁶ “Naively understood,” because acts performed under different descriptions are not, in some sense, the “same act” at all (Davidson 1980). But this external aspect of the act is all that outside observers can confidently monitor; and if, as will be argued, agents are often anxious to keep their options open as to how to describe an act, even to themselves, that is all *anyone* will ever have to work with. Philosophers anxious to describe an action in terms that the actor would recognize as his own must not foreclose arbitrarily the possibility that he is himself insistently ambivalent as regards act-descriptions, even in his own mind.

⁷ Practice rules of the sort here in view have been adopted deliberately and are therefore immune to the worries of Ryle (1949, pp. 110–13).

In addition to those grounds for expecting there to be a multiplicity of convergent reasons in general, there is a special ground for expecting this to be particularly true in sociopolitical affairs. This connects up with what used to be called the "functional requisites of a stable society." That phrase puts the point too strongly, and introduces all sorts of awkward questions about agency as well. In a gentler formulation, however, the point is well nigh indisputable. There are all sorts of advantages to living in a society which is arranged in such a way that everyone finds it in his private interest to do what is in the public interest.

How, exactly, that coincidence of interests comes about is an open question. Sometimes it is through a happy accident. Other times it is through the invisible hand of the market, or other analogous institutions based on reciprocal, conditional cooperation. Still other times it is through the intentional intervention of social engineers "designing institutions for knaves." For present purposes, the precise mechanism does not matter.

The point is merely that, in any well-ordered society, affairs are arranged in such a way that acts of altruism (along with other forms of morally laudable behavior) will usually, if not quite invariably, pay the person who performs them. In such cases, the same convergence of public and private interests that underwrites the stability and smooth functioning of social systems also guarantees that people will ordinarily have multiple, converging reasons – some egoistic, some altruistic – for performing the same action.⁸

IV

Wherever there are, in any of these ways, many converging reasons for an agent to perform one and the same action, there are no particularly good grounds for that agent himself to bother pinning down his own "true" motives, that is, the ones "really" guiding his behavior.⁹ In such circumstances as these, concern with motives is primarily an external concern rather than an internal one. It is the sort of thing

⁸ The agent might be led to say, with MacIntyre (1967, p. 466), "it is not . . . that I have two separate motives, self-interest and benevolence, for doing the same action. I have one motive, a desire to live in a certain way, which cannot be characterized as a desire for my good rather than that of others."

⁹ For empirical evidence, see Nisbett and Wilson 1977. Furthermore, most actions are complexes or sequences across a protracted period of time; and at different moments different ones of an agent's multiple, converging reasons for action will weigh more heavily with him. Thus, even if he could pin down which was his "one true motive" at each moment, there remains the question as to which in this long sequence of shifting motives should be called "the" motive underlying the action-complex as a whole.

that other people worry about – and have good grounds for worrying about. Their worries are not, however, ones that the actor himself has any particularly strong *pragmatic* grounds for sharing. That is established in Section IVA. The agent may well have strong *moralistic* grounds for inquiring into his motives, nonetheless. But, as I shall show in Section IVB, such inquiries will prove fruitless – the answers inherently inconclusive and unreliable – if that is his *only* reason for inquiring.

A

We characteristically demand to know *why* an agent did what he did under either of two circumstances. One is when we can see *no* reason to do it; the other is when we can see *too many*. In the former case, we are demanding that the agent give us “one good reason” for behaving as he did. In the latter case, we are demanding much the same. Where we can see that he might have done something for any number of very different reasons, we once again want to know which one (or ones) of those converging reasons were *his* reasons – which were foremost in his mind when choosing that action.

The first purpose such inquiries might serve is a pragmatic one. We want to anticipate other people’s future behavior in order to plan our own so as to pursue more effectively our own ends, whatever they may be. Knowing people’s motives is crucial in this connection, because motives provide clues to understanding action – past, present and future. Coming to understand a person’s motives for acting as he did enables us to explain his past behavior, and to do so in such a way that allows us to predict his future behavior (assuming, of course, that he will be similarly motivated in similar future situations).

Notice, however, that this concern weighs far more heavily with other people than with the actor himself. From a pragmatic point of view, it is far more important to us to be able to isolate the motives underlying other people’s behavior than it is to be able to isolate the motives underlying our own. That is simply because predicting other people’s future behavior is much more important to us than predicting our own. Other people’s behavior forms the backdrop against which we will have to act. It constrains our choice of actions to some greater or lesser extent; or it facilitates our choices to some greater or lesser extent. We want to predict other people’s behavior in order to know what choices will be effectively available to us, and on what terms.

From our own point of view, others’ actions are to be *predicted*; our own are to be *chosen*. If others act in ways out of keeping with the pattern formed by their past behavior, our own plans (insofar as they are predicted on a presumption that those patterns will persist) will

have been spoiled. Insofar as we act in ways out of keeping with the pattern formed by our own past behavior, our own plans will merely be said to have changed. There is no reason – no direct, internal reason, at least – why our plans should not change, or for us to regret such a change if it does occur.

That is not to say that an agent has absolutely no reason for wanting to be able to predict his own future behavior. For one thing, he has some indirect interest in doing so, derived in various ways from the interest that others take in it. For another, an agent whose preferences are set to change will want to know when and how, either with a view to choosing plans whose future payoffs will be agreeable to his future self (presuming he happens to take an interest in his future self's satisfactions) or, minimally, with a view to using up resources that will cease to be useful to him once his plans have changed.

None of those sorts of goals, however, strictly requires clear knowledge of your own present motives. If you want to give others firm guarantees of your future behavior, the way to do that is not to parade your present motives before them; it is rather to superimpose new and stronger motives for behaving as you promise (e.g., by signing binding contracts which are costly to break). Similarly, if you want to know when and how your preferences are going to change, looking within yourself and your present desires for clues is probably not the best way to find out. You would be better advised to look instead outside yourself, to the external forces that will shape your future desires.

Even if such considerations do apply to the agent himself, they do not apply to him anywhere nearly as strongly as they do to others. He has less of an interest in pinning down his own motives than others do, because he has less of an interest in predicting his own future behavior than they do. When confronted with an inconsistency between his previous plans and his present actions, the agent himself always has the option of declaring, "I changed my mind." From his own perspective, that successfully transforms what would have counted as a cost in terms of his previous plans into a benefit in terms of his new ones. The fact that an agent's altered actions effectively serve his new goals is, however, of little comfort to others whose plans for pursuing their own goals were predicated upon predictions of that agent's behavior which have now proved false.

Much the same can be said of a person who has not actually changed his mind, but merely does not know his mind. This is the case of someone who persistently makes incompatible choices, undoing with one action what he has just done with the last. It appears that he is moved by inconsistent, poorly ordered motives/preferences.

And it therefore might appear that he has much the same sort of an interest in understanding and sorting out his own motives, for his own purposes, as others have in understanding his motives for theirs.

But that argument moves too quickly. It too readily tars motivational inconsistency with a brush best reserved for cognitive inconsistency alone. Arguments from consistency are conclusive, perhaps, as regards questions of truth. If two beliefs can be shown to be inconsistent, then (according to all standard logics, anyway) one of them simply has to be abandoned as false. In the realm of desire, however, truth is not at issue. There, consistency is a very different sort of virtue, and inconsistency carries a very different kind of penalty."

To be sure, two inconsistent desires cannot be simultaneously satisfied. That is a powerful (pragmatic) reason for not pursuing them *simultaneously*. But that is not necessarily an argument for not pursuing them *separately* in different spheres of one's life, pigeonholing the paradoxes, as Wittgenstein would say. Nor is it an argument for not pursuing them *sequentially*. If the desires really are incompatible, then of course it will remain true that in pursuing the one you will be undoing whatever you have already accomplished in pursuit of the other. But it is simply incorrect to say that you are necessarily worse off at the end of this process, if you truly harbor both desires. You may be going around in circles, but saying that you harbor both inconsistent desires may be merely to say that you enjoy going around in that particular circle." The upshot may be much the same as with the agent with changing preferences – an agent can derive satisfaction from the (counterproductive) pursuit of his inconsistent desires; and this satisfaction is of a sort that is necessarily unavailable to others who are affected by his inconsistent choices.

In conclusion, others need to know our motives because they need to know how to get from us whatever it is they want in any particular circumstance. We, for our own part, need only know what we want

" B. Williams 1973b, chaps. 11 and 12. Nozick 1981, pp. 405–9.

" The classic proof that intransitivity is positively irrational is that someone with such preferences can be induced to make "Dutch book" against himself. If he prefers A to B to C to A, then he would be willing to pay some positive sum of money to make each move, going around in a never-ending circle until he was broke (Davidson and Suppes 1957, p. 2; Raiffa 1968, p. 78). That is supposed to be a *reductio*, but someone who really did prefer A to B to C to A might reckon it money well spent, enjoying as he does each ride around the merry-go-round.

In any case, it is unclear how understanding the *motives* underlying your own intransitive choices would have helped you avoid the disastrous consequences of pursuing them to your ruin. Mostly what is needed for that is simply knowing *that* you have intransitive preferences and what follows from that fact; knowing *why* you do so is neither necessary nor even particularly helpful in avoiding their consequences.

in any particular circumstance. We need to know others' motives, for the same reason they need to know ours. But there is no pragmatic need for us to inquire deeply into our own motives, just so long as they are not so unclear or so inconsistent as to stymie choice itself.¹²

B

We might have a second purpose for inquiring into motives. This is a moralistic one. We want to know why a person did what he did, because we want to decide what sort of a person he is. Consequentialists might regard this inquiry as just a variant on the other. For them, all concern with people's motives and character is reducible back to a desire for indicators of their future behavior. But for deontologic moralists in the Kantian tradition, or the Judeo-Christian tradition more generally, questions of motives and character retain some considerable independent interest. For them, the quality of an act depends crucially on the quality of the will from which it proceeds. For purposes of moral evaluation, conducted in this mode, at least, we must know the agent's intentions as well as his actions (past, present or future).¹³

In Section IVA, I argued that knowing someone's true motives is pragmatically more important for others than for the agent himself. Moralistically, that is decidedly not true. An agent has a direct, immediate interest in his own moral standing that is at least as strong as (and arguably far stronger than) anyone else's interest in it. The question to be raised here, then, is not whether an agent has an interest in inquiring into his own motives. For moralistic purposes, surely he does. The question is instead whether such inquiries are likely to prove fruitful, in cases of multiple, converging motives, if that moralistic purpose is the *only* purpose for which these inquiries are being conducted.

The situation in view, remember, is one in which there are multiple, converging reasons for performing one and the same action. The task before the agent is to determine, for certain moralistic purposes, which of those reasons weighs most heavily with him. As argued in Section

¹² Simultaneously preferring A to B and B to A, for example, would render choice impossible; so too would an incomplete preference ordering, at least insofar as unranked options are concerned (Goodin 1976, chap. 2).

¹³ Intentions are not strictly equivalent to motives, as the notion of *mens rea* in the criminal law makes clear (Hart 1968, chap. 5). What matters there is simply that the criminal performed the act in question, fully intending its natural consequences – *why* he sought those consequences is neither here nor there. Although judgments of legal guilt are substantially independent of such motivational inquiries, judgments of blameworthiness or excusability – and hence of appropriate punishment – typically do turn on them (Fletcher 1979; Brandt 1985).

IVA, however, there is no pragmatic need for him to settle the issue. So the agent is addressing this moral question about motivations unconstrained (or minimally constrained) by pragmatic, objective reality. He can give any of a great number of different answers to the moralistic question, without having had to alter his behavior in the slightest.

What are the results of this inquiry worth, when there are so few external checks or constraints on them? For the conclusions of this moralistic inquiry into motives to carry any real significance, it is crucial that the agent should be reporting truly what relative weights he really does (or did) assign to different considerations. But neither he nor we have any way of knowing whether what he reports as his motives really are his true motives; after all, the results of this inquiry would not have made any pragmatic difference to what he decided to do, in the end. With nothing else at stake in the inquiry *except* his moral standing, the agent doing the accounting would enjoy enormous latitude to cook the books at will. Where that is so, nothing much can follow from the accounting exercise.

The point is not just one about veracity in reporting one's own motives. The fear is not so much that the agent will lie, but that without any reality check neither he nor we will have any way of telling what the truth of the matter really is. Nor is the worry that he will necessarily cook the books in his own favor, attributing to himself nobler motives than he in fact harbors. He may do just the opposite, engaging in moral self-debasement and attributing to himself less noble motives to himself than are really at work. The point is merely that there is no way of checking the books, and that fact alone makes them worthless accounts for the purposes at hand.

An agent, anxious to inspect his own motives to discover what sort of a person he is, would be better advised to turn his attention to other sorts of situations entirely. Instead of examining cases where many diverse reasons for action converge, and trying to decide which of the several motives (all of which point in the same direction) really guided him, the agent should instead look to cases where reasons for action sharply diverge. Where different motives would point the actor in different directions, he can readily determine which motive really guided his behavior merely by observing what, in the end, he did. Such crucial test cases allowing for the unambiguous differentiation of motives will be rare. That is the burden of the preceding arguments; and that conclusion is confirmed by the common experience of different biographers, all intimately acquainted with the details of some subject's life, but nonetheless irreconcilably at odds in their final judgments as to what sort of a person he really was. Still, it is to those rare test cases, rather than to the more common case of multiple and

converging reasons, to which we must look in order to decide what sort of a person one really is. (Likewise, it is to those rare cases of "many reasons for doing the wrong thing, and only a single morally noble one for doing the right thing" to which we should look in our search for moral exemplars.) In the more common cases of converging reasons for any given action, we might (at most) merely want to check to make sure that the moral reason is at least present among all the others. It is futile to try to decide, under those circumstances, whether it is the controlling motive.

V

The burden of the foregoing argument is that, where there are many different reasons for an agent to perform one and the same action, the agent himself has *no good reason* to try to disentangle his own true motives. That established, I next want to make a stronger and more positive claim. There might often be good reasons for him to do exactly the opposite. That is to say, there might often be good reasons for an agent positively to *avoid* pinning down his own true motives too tightly, certainly in public and perhaps even within his own mind.

Such obfuscation might carry advantages of a perfectly general sort. Here, however, I shall concentrate more narrowly on the motives of altruism and egoism, and on the strategic advantages that can often accrue from fudging the question of which has motivated any particular piece of behavior. I shall identify three ways in which such strategic advantages might accrue, without presuming them to be either exhaustive or necessarily mutually exclusive.

1. Notice, first, that altruism might be seen as entailing a kind of commitment. When you declare yourself to be behaving altruistically this time, without further elaboration as to what peculiar brand of altruism you might be practicing, you implicitly commit yourself to doing similarly in at least some similar circumstances in the future. But you may not want to make that sort of a commitment. You may regard benevolence as a kind of Kantian imperfect duty. Giving alms to some beggars should suffice; just because you give alms to some beggars, that does not mean that you are willing to commit yourself to giving alms to all (even to any other) beggars. You may want to be altruistic with respect to some particular individuals – your family, friends, compatriots – without committing yourself to being equally altruistic toward the world at large, even though admitting that you can offer no good grounds for limiting your favors thusly. In those circumstances, too, you are strategically well advised to resist describing your actions in terms anything like as general as "altruism." Absent

some plausible principle to justify your limited loyalties, that description would stand in serious danger of committing you to more than you would wish.

2. Just as you obfuscate as regards your private motives so as to avoid unnecessarily committing yourself, so too might you want to obfuscate as regards the collective intentions underlying some joint endeavor so as to avoid losing supporters who would not want to commit themselves unnecessarily to similar future action. Different people join in any particular piece of collective action (enacting a particular piece of legislation, overthrowing a particular dictator) for a variety of different reasons.¹⁴ Vague and contradictory statements of collective intent not only reflect the tensions contained within such marriages of convenience; they are also option-preserving elements which are crucial in helping to consummate such marriages in the first place.¹⁵

3. Sometimes a person is prepared to be altruistic, but only in his dealings with other altruists. It would be to your advantage to be treated altruistically by this person. But to do so, you will have to feign more altruism than you really feel – or at least you will have to give sufficiently unclear signals that a generously disposed conditional altruist will be prepared to interpret them as signs of altruism on your part. It would be a smart move for you to feign altruism or to obfuscate in this way, just so long as the gain that you would secure by being treated altruistically by others exceeds the loss that you would suffer by treating them less ruthlessly than you would otherwise have done. An altruist who unilaterally declared himself to be playing an Assurance Game might be able to elicit a fair bit of “as if” altruism from others, as well.¹⁶

Taken together, these three propositions suggest that it might often be in an agent's strategic interests to obfuscate as regards his true motives. In saying “strategic interests” here, I do not necessarily mean to imply that their true motives are invariably egoistic. An altruist, too, would find it in his strategic interests – that is, it would promote the happiness of others better – if he were to obfuscate as regards his motives for passing money to needy others. Recipients feel better getting money that they can plausibly regard as a “loan” rather than as a gift of pure charity, for example.

¹⁴ Brennan (1973) offers the delightfully perverse example of both altruistic and envious people agreeing, each for his own very different reasons, to a redistribution of income from the rich to the poor – the altruist acting out of sympathy for the poor, the envious out of hatred for the rich.

¹⁵ MacCallum 1966.

¹⁶ Sen 1967.

VI

The practical implication of all of this is simply that it is often a mistake to insist not only that the right thing be done, but that it also be done for the right reason. By forcing the motivational issue in this way, we will often succeed only in reducing the frequency with which the right thing gets done at all. Furthermore, by forcing the issue in this way, we may even be putting off some people who would have done the right thing for the right reason privately, but who refuse (for one of the strategic reasons mentioned) to do so when an unambiguous, public revelation of motives is required. The conclusion is that we ought to abandon inquiries into motives altogether, at least for that large class of cases characterized by multiple, converging reasons for action.

There may still remain some practical problems of deciding how to respond to some particular piece of behavior without knowing for certain what its motives were. Here I suggest we simply borrow a model devised by U.S. courts for dealing with policies that might have been motivated by racial or sexual antagonisms. Although such discrimination by government agencies is of course illegal, a showing of discriminatory intent is not, in and of itself, sufficient to settle the issue; that merely shifts the burden of proof onto the agency, which then has to prove that the policy (however discriminatory in its intent) actually serves some other legitimate purpose. If the agency can in this way show that it is a good policy, then the courts will not overturn it even if the agency adopted it for "bad" reasons.¹⁷

This is an approach which my analysis would encourage moralists to adopt, likewise. If there were enough good reasons for an action, then it does not matter how many bad ones there were for it as well; nor does it matter which of those several reasons really guided the agent's choices. Certainly we should arrange our social affairs in such a way that people can act on noble motives if they so desire, and doing so may require fairly radical restructuring of some aspects of society (e.g., the economy). The mistake comes merely from insisting that people must act purely from noble motives.¹⁸

¹⁷ Ely 1970, pp. 1282ff. R. Bennett 1979.

¹⁸ For helpful discussions of these issues, I am grateful to Lincoln Allison, Alistair Edwards, Vinit Haksar, David Hunter, David Miller, Max Neutze, Philip Pettit, Raymond Plant, Andrew Reeve, Hillel Steiner, Hugh Stretton, Richard Sylvan, Michael Taylor, Jeremy Waldron, Hugh Ward and Albert Weale.

Government house utilitarianism

Critics of utilitarianism invite us to contemplate alternatives to utilitarianism. When doing so, however, we ought also to consider alternatives within utilitarianism. Not all utilitarianisms are of a cloth, and some of the criticisms that might be devastating against one variant might not stick against another. Depending upon what, precisely, it is that bothers us about utilitarianism, utilitarianism itself might provide an answer.

The ordinary way of denominating varieties of utilitarianism, surveyed in Chapter 1, differentiates them according to what the felicific calculus is to be used to choose. Thus, act-utilitarianism has us choosing actions, one by one, according to the calculus of pleasures and pains. Rule-utilitarianism has us choosing a rule of conduct that will, insofar as it is followed, maximize utility, and that rule then dictates our choice of actions.¹ Motive utilitarianism has the utility calculus being used to select motives and dispositions according to their general felicific effects, and those motives and dispositions then dictate our choices of actions.

The distinction I shall here propose works along a dimension orthogonal to that one.² Instead of differentiating utilitarianisms on the

¹ The phrase "insofar as it is followed" here deliberately straddles two possible alternative formulations: (1) "choose that rule which, if followed by everyone, would maximize utility" and (2) "choose that rule which, given what others will do, maximizes utility if chosen to guide your actions." The latter is clearly the more correct formulation of the principle when used by individual choosers (Harrod 1936, pp. 151-2; Regan 1980). The former, or some variation on it, is arguably the more defensible when the principle is used by those responsible for coordinating the actions of a large group of agents.

² "Orthogonal," because for various reasons offered later public officials are precluded from the direct pursuit of utility and adopt indirect, rule-based strategies instead. Thus, many of the advantages of the form of utilitarianism I here rec-

basis of what they are used to choose, I suggest doing so on the basis of *who* is supposed to use the utilitarian calculus to make choices. Implicitly, contemporary discussions of varieties of utilitarianism are all standardly addressed, first and foremost, to individuals acting in their personal capacities and making choices which, while they may affect others as well, principally affect the choosers' own lives. Implicitly, public officials, choosing in their official capacities and on behalf of others, are expected to address their choices in exactly the same fashion, bringing exactly the same considerations to bear as would private individuals.¹

That, I submit, is a fallacy. It does matter who is using the utilitarian calculus, in what circumstances and for what purposes. Using the felicific calculus for micro-level purposes of guiding individuals' choices of personal conduct is altogether different from using it for macro-level purposes of guiding public officials' choices of general social policy. A different menu of options – in some respects greater, in others, less, but in any case different – is available to public and private choosers.

Those differences are such as to neutralize, in the public sphere, most of the objections standardly lodged against utilitarianism in the private sphere. True though such complaints may be as applied to utilitarianism as a standard of personal conduct, they are irrelevant (or anyway much less problematic) as applied to utilitarianism as a standard of public policy. Or so I shall argue.

Dubbing the form of utilitarianism I advocate "government house utilitarianism" seems apt, since it does after all concern the formulation and implementation of public policy. It also runs the risk of unfortunate echoes, of course. That is after all a term usually applied, usually derisively, to the closing chapters in Sidgwick's *Methods of Ethics*, where he recommends that enlightened (implicitly, colonial) rulers govern according to utilitarian principles that are not necessar-

commend flow from the advantages of indirect rule-utilitarianism more generally. Still, there are peculiar advantages to be derived from making public policy in these ways, and there are peculiar constraints on public policy-makers obliging them to behave in this fashion. Hence, though this argument for "government house utilitarianism" is related to the case for indirect rule-utilitarianism more generally, it is not wholly reducible to it.

¹ In contemporary discussions, it is surprisingly rare to find any explicit acknowledgment that utilitarianism can serve both functions, guiding public as well as personal choices. When the distinction between the two functions is noticed at all, the former is almost invariably given primacy – at least in discussions of moralists and even philosophers, if not necessarily in those of economists. See B. Williams 1973a, pp. 138–40, and Sen and Williams 1982, pp. 1–2; cf. Hardin 1986; 1986 and Sen and B. Williams 1982, chaps. 2–4, 9 and 10.

ily (and perhaps necessarily not) accessible to those subject to their rule.⁴ I do not wish to commit myself in advance to this intuitively unpalatable conclusion – though of course as a utilitarian I cannot commit myself firmly against it, either, ahead of a calculation of consequences. What I can say, *ex ante* of any such calculation, is simply that it is not this aspect of government house utilitarianism that attracts me, and it is not on that basis that I propose to commend it to others. If this is indeed a consequence, it is a (preanalytically) unwelcome one.

One final prefatory note. In urging utilitarianism as a guide for policy choice, I am being faithful to the tradition. Whatever contemporary writers might say, “The fathers of utilitarianism thought of it principally as a system of social and political decision, as offering a criterion and basis of judgment for legislators and administrators”; and “this is recognizably a different matter from utilitarianism as a system of personal morality.”⁵ Bentham’s *Introduction* was to the principles of morals *and* legislation, after all; and to judge from Bowring’s collection of his *Works*, he spent the vast majority of his time advising on matters of constitutional and penal law reform and other sundry topics in public policy and administration. And so too with his most immediate successors.⁶

I. THE PECULIARITIES OF PUBLIC-POLICY MAKING

My larger argument turns on the proposition that there is something special about the situation of public officials that makes utilitarianism more plausible for them (or, more precisely, makes them adopt a form of utilitarianism that we would find more acceptable) than private individuals. Before proceeding with that larger argument, I must therefore say what it is that is so special about public officials and their situations that makes it both more necessary and more desirable for them to adopt a more credible form of utilitarianism.

A. The argument from necessity

Consider, first, the argument from necessity. Public officials are obliged to make their choices under uncertainty, and uncertainty of a

⁴ Sidgwick [1874] 1907, bk. 4, chap. 5, secs. 2–3. Cf. B. Williams 1973a, pp. 138–40, and Sen and Williams 1982, p. 16.

⁵ B. Williams 1973a, p. 135.

⁶ Bentham [1789] 1970; 1843. Among Bentham’s followers, John Austin clearly used the utilitarian calculus for judging rules and institutions (Rawls 1955), and J. S. Mill at least arguably did so (Urmson 1953; cf. Mabbott 1956).

very special sort at that. All choices – public and private alike – are made under some degree of uncertainty, of course. But in the nature of things, private individuals will usually have more complete information on the peculiarities of their own circumstances and on the ramifications that alternative possible choices might have for them. Public officials, in contrast, are relatively poorly informed as to the effects that their choices will have on individuals, one by one. What they typically *do* know are generalities: averages and aggregates. They know what will happen most often to most people as a result of their various possible choices. But that is all.

That is enough to allow public policy-makers to use the utilitarian calculus – if they want to use it at all – to choose general rules of conduct. Knowing aggregates and averages, they can proceed to calculate the utility payoffs from adopting each alternative possible general rule. But they cannot be sure what the payoff will be to any given individual or on any particular occasion. Their knowledge of generalities, aggregates and averages is just not sufficiently fine-grained for that.

For an example, consider the case of compulsory seat belt legislation. Policy-makers can say with some confidence that, on aggregate, more lives would be saved than lost if all automobile drivers and passengers were required to wear seat belts. As always, that aggregate conceals the fact that some gain while others lose. Some people would be trapped by seat belts in fiery crashes who would otherwise have been thrown to safety by the force of the impact, after all. The point is that policy-makers, contemplating seat belt legislation, have no way of knowing who those individuals are, exactly, or on what occasions, exactly, that might occur. All they can know is that, on aggregate, far fewer people would be saved than killed by being thrown clear of their cars upon impact.

Furthermore, the argument from necessity would continue, the instruments available to public policy-makers are relatively blunt. They can influence general tendencies, making rather more people behave in certain sorts of ways rather more often. But perfect compliance is unrealistic. And (building on the previous point) not knowing particular circumstances of particular individuals, rules and regulations must necessarily be relatively general in form. They must treat more people more nearly alike than ideally they should, had we perfect information.

The combined effect of these two factors is to preclude public policy-makers from fine-tuning policies very well at all. They must, of necessity, deal with people in aggregate, imposing upon them rules that are general in form. Nothing in any of this necessarily forces them to be utilitarian in their public policy-making, of course. What it does

do, however, is force them – if they are inclined to be utilitarian at all – away from direct (act) utilitarianism and toward something more indirect, like rule-utilitarianism.⁷ The circumstances surrounding the selection and implementation of public policies simply do not permit the more precise calculations required by any decision rule more tailored to peculiarities of individuals or situations.

B. The argument from desirability

The argument from desirability picks up where the argument from necessity leaves off. It is a familiar principle of natural justice that people ought to be governed according to laws that are general in form, rather than through particularized edicts applying to small subsets of the population alone (the limiting case of which is the much maligned “bill of attainder”). Of course, a utilitarian is in no position to help himself to principles of justice willy-nilly, without offering some broadly utilitarian account of the wisdom of those principles.⁸ However, such an account can, I think, be given.

The more high-minded version is this. If laws have to be general in form, and apply to everyone alike, then we can make some pretty shrewd guesses as to what sorts of future laws might be enacted; and we can plan our own affairs accordingly. If particularized rules (or substantial discretions in applying the rules) are permitted, then anyone and everyone might be made an exception to the general rule. Under such circumstances, no one can know for sure what will be required of him in the future.⁹ Yet there are substantial utilitarian gains – both to the individuals themselves, and to others whose own plans depend for their success upon the actions of those individuals – from being able to enter into long-term commitments in some confidence that they will indeed be carried out.¹⁰ From all that, it follows that there are substantial utility gains from requiring that laws be

⁷ Or like motive-utilitarianism, perhaps: but while policy-makers certainly can engage in mass character-building to a limited extent, they can do considerably more to legislate rules than motives or dispositions. On indirect consequentialism more generally, see R. Brandt 1988.

⁸ Alternatively, he can accept those principles as “moral side-constraints” on utilitarian maximization, as did arguably all the classical utilitarians (Rawls 1955, p. 9). There is no reason to suppose that such side constraints would be structurally incompatible with basic utilitarian maximization – all maximization occurs under constraints of some sort or another, and the constraints have to be awfully tight before the element of maximization loses its interest (Elster 1979b, pp. 113–14).

⁹ It would be small consolation that legislators or administrators make those decisions according to an act-utilitarian calculus, for its dictates are just as unpredictable as the whims of the majority.

¹⁰ Hume 1739, bk. 3, pt. 2, sec. 2. Hodgson 1967. Harsanyi 1977b.

relatively general in their form and hence relatively predictable in their content.

Another way of arguing for the desirability of that practice, still thoroughly utilitarian in form, is this. Enforcement costs are in utilitarian terms a deadweight loss to be minimized insofar as possible. One way to minimize such costs is through the self-regulation of people's conduct. If people can be brought to internalize social norms, adopting them as their own and shaping their conduct accordingly, there would be no need for expensive enforcement measures, with obvious utilitarian advantages. But for principles of conduct to be easily internalized in this way, they must be few in number and general in form. If the idea is to let people govern their own conduct in line with rules, then they must be able, first, to learn and recall all the relevant rules when the occasion demands; and, second, to apply the rules to particular situations for themselves, without the aid of authoritative guidance in each instance." All of that is easier, and hence the utilitarian payoffs higher, the less numerous and less complex the rules are.

Whereas the classic argument from justice is that it is "only fair" that people be governed according to general rules, the utilitarian argument from desirability is that it is "only prudent" to do so. In that way, people can largely anticipate what the rules will require of them, and apply the rules for themselves without expensive social enforcement.

II. CRITICISMS OF UTILITARIANISM BLUNTED

Thus I am left concluding that public policy-makers, given their special circumstances, both ought and in any case must issue orders that are general in form. That, in turn, serves to blunt many of the criticisms of utilitarianism that are rightly lodged against its use as a code of personal, private conduct, as I shall now attempt to show.

A. Utilitarianism asks too much

One familiar argument against utilitarianism is that it asks *too much*¹⁴ of us. Under it, we must be always and ever prepared to engage in

¹⁴ As argued, variously, by Rawls (1955, pp. 23-4), Hart (1961, p. 127), R. Brandt (1963, p. 125), Hare (1981, pp. 35-6) and Goodin (1982b, chap. 4).

¹⁵ The "too much/too little" formulation of this and the next section derives from Kagan (1987, p. 644), who argues that these are the "central objections" to utilitarianism: "There may be other 'problems' with utilitarianism, but permitting the impermissible and demanding the nonobligatory - these are the *sins* of utilitarianism," in the eyes of its deontologist critics.

good works – however much they demand of us, however disruptive they may be to our own lives – whenever a utilitarian calculus shows that others stand to gain more than we would lose. Such constant willingness to engage in sacrifice is asking awfully much of people, though. Perhaps we should all be like that, but in practice it is more than most of us can muster. It is more than is realistic – or hence reasonable – to expect of most of us. That may be the stuff of which saints and heroes are made. For more ordinary mortals, however, there should be some “limits of obligation,” some “cut-off point” beyond which performances should be counted good but utterly supererogatory, above and beyond the call of duty.¹¹

The sorts of circumstances in which this concern seems most legitimate are those sketched in Peter Singer’s frankly utilitarian account of “Famine, Affluence and Morality.”¹² Hedge his argument though he will, Singer cannot avoid the specter of his First World readers being required to provide aid to the starving of the Third World at such a level, and with such frequency, that they will themselves be reduced to the status of the “almost starving.” That might be right morally. Still, most readers will find it a conclusion that is pretty hard to stomach.

Most of those problems arise merely through a lack of coordination, though. Why should I impoverish myself to feed the world, while others situated similarly to me are not lifting a finger to help? Again, perhaps the right answer (rankle though it may) really is that you should do so.¹³ Or, more precisely, perhaps you should cooperate with as many others as are willing to cooperate with you, to do as much good as you (collectively) can.¹⁴ But, again, it seems to be asking an awful lot.

Such embarrassments for the utilitarian arise, at root, from addressing moral injunctions to individuals alone. Individuals must take the actions of other individuals more or less as given. They may try to persuade others to join in their good works; they may even try to shame others into joining; but they may not compel them to join. Because of that, the conscientious utilitarian does indeed run a very real risk of being required to do too much, in some sense or another. Insofar as he will be required to make up for others’ moral delicts, he certainly will end up doing more than his fair share; and depending

¹¹ Urmson 1958. Fishkin 1982. Heyd 1982. It is an open question whether any properly principle-based (rather than merely intuition-driven) deontology would not, in effect, demand as much of people, as Kagan (1987, pp. 649ff.) rightly remarks in reflecting on writers like Donagan (1968; 1977).

¹² Singer 1972.

¹³ Singer 1972; 1979, chap. 8.

¹⁴ Harrod 1936, pp. 151–2. Regan 1980.

upon just how much that is, he may well be required by utilitarianism to do more than psychologically he can bring himself to do. It is precisely because he, *qua* individual chooser, cannot choose for others to do their duty that he gets stuck doing double duty – his own, and theirs as well.

That is where public officials are in a different position. They can choose how much we, collectively, will give to the Third World; and those choices can be made binding on the rest of us. Certainly by utilitarian standards, and probably by most others as well, public officials may legitimately compel us all to play our required parts in coordination schemes for the discharge of our shared moral duties.¹⁷

Given this fact, it is much less likely that utilitarian public policies, enforced upon all alike, will “ask too much” of any one of us. Since others will be required to do their share, too, the demand laid upon any one person will be considerably less than it would have been had the burden been shared only among the relatively few who would have contributed voluntarily.¹⁸

Even with coordination to guarantee that everyone does his share, utilitarian calculations might nonetheless require a very few very rich people or societies to make great sacrifices in order to improve (perhaps only slightly) the welfare of a great many very poor ones. But what are we to make of the case where suffering runs so deeply and so widely as to be alleviated only by sacrifices that are great, even when equitably distributed among all the rich of the world? It should be regarded as a tragic situation to be sure. But why should it be regarded as a tragedy for the poor alone? Why suppose that the rich should be able to escape the effects of the tragedy?

Tragedies call forth heroes, and not just in the sense of giving people's latent heroism an opportunity to manifest itself. Sometimes, tragedies make sacrifices that would, under ordinary circumstances, count as heroic amount to “the least you could do.” In tragedies, we expect heroes. Or, less paradoxically, in tragic situations we expect people to behave in ways that, absent those exceptional circumstances, would hardly be required and might even be regarded as heroic. Exceptional circumstances demand exceptional deeds.

¹⁷ As I have argued in Chapter 2.

¹⁸ There is, of course, the international analogue of the “heroic sacrifice” problem. In a world of independent states, a superconscientious state might by similar utilitarian reckoning be required to make up for the delicts of less conscientious ones. The solution here is much as before – just locate Government House in UN Plaza. Just as the enforcement of a coordination scheme domestically can be justified in terms of the legitimacy of compelling people to play their required parts in schemes for the discharge of shared duties, so too can enforcement of international schemes be similarly grounded.

Beyond all that, there is the further question of what is the equivalent, at the level of society as a whole, of "asking too much" of a person. At the level of the individual, the notion is cashed out in terms of psychological capacities. There are some things that most people, given their psychological makeup, just cannot bear to do. But societies do not, literally, have minds, psyches or, hence, limited psychological capacities.

It is of course true of us as a society, just as it is of us as individuals, that there are some sacrifices that are so great that we cannot make them and still maintain whatever it is that makes us who and what we are. But that, in itself, is not what got the "saints and heroes" argument going. The crucial move in that argument was the next one: that real people cannot psychologically bear to see their identity wiped out in that way. True, similarly large sacrifices might similarly compromise a society's identity. But it is nonsense to say that that is something that society, literally, cannot psychologically bear. In the end, it must always come down to a matter of asking too much of people *in* societies. Asking them to sacrifice important aspects of their shared life in such ways might be more than individuals can psychologically bear.

The larger question is whether such statements – about an individual's own projects or socially shared ones, either – are to count as claims or confessions. There may well be people who are psychologically so attached to their claret club that they cannot bear to think of its being disbanded; to do so would undermine their sense of self and, with it, their very capacity for moral agency. But to say that they cannot bear to contemplate abandoning their luxuries so that others might be given the necessities of life is to say that their capacity for moral agency was pretty meager all along. How to treat subnormals is always a tricky question. The standard answer, though, is surely to humor them but to deprive them of any power to harm themselves or others. Those who confess to their limited moral agency in such ways may be similarly deserving of the moral equivalent of a padded cell or sheltered workshop.

B. Utilitarianism asks too little

Sometimes it is said that utilitarianism asks too little of us.¹⁹ For it, nothing is right or wrong, morally prohibited or morally required, in any and all circumstances. There is nothing necessarily sacrosanct about people's rights or liberties or integrity; there is nothing neces-

¹⁹ B. Williams' (1973a, p. 137) "Critique of Utilitarianism" is largely centered on "how little of the world's moral luggage it is prepared to pick up."

sarily sacrilegious about violating the Ten Commandments. It all just depends upon how the numbers come up in the giant utilitarian adding machine. There are none of the "moral side constraints" that Nozick in his way and Donagan in his would wish to impose upon consequentialist calculations.²⁰

It may well happen, at the level of personal ethics, that utilitarian calculations would require us to do things that would violate the sorts of side-constraints that deontologists would impose. That is the force of Bernard Williams's famous examples. Jim might have to violate the right to life of one Indian to slake Pedro's taste for the blood of the other nineteen. Or it might be better for a young chemist, George, who is himself opposed to the manufacture of war material, to go to work for a napalm manufacturer. His own integrity would be compromised in so doing, of course, but that would prevent the post from going to another researcher whose search for a better sort of napalm would be far more zealous than George's own.²¹

Again, I would like to leave it as an open question what is the right thing to do in those circumstances. If the circumstances were really as described, then they are very different indeed from those around which our ordinary intuitions of right and wrong have been framed; and, counterintuitive as it seems, it may well be right for us to do precisely what utilitarians recommend in such strange cases.²² There is, after all, a certain moral preciousness involved in arguments about people's "integrity" and "clean hands." To paraphrase Brian Barry, if I were one of the nineteen Indians Jim could have saved, I would not think much of this moral dandy who prates on about his integrity while people die needlessly.²³ So even at the personal level, it may not be so obviously wrong to do as utilitarians recommend.

My main argument, though, is that at the level of social policy the problem usually does not even arise. When promulgating policies, public officials must respond to typical conditions and common circumstances. Policies, by their nature, cannot be case-by-case affairs. In choosing general rules to govern a wide range of circumstances, it is extraordinarily unlikely that the greatest happiness can ever be realized by systematically violating people's rights, liberties or integrity – or even, come to that, by systematically contravening the Ten Commandments. The rules that maximize utility over the long haul

²⁰ Donagan 1968; 1977. Nozick 1974. Kagan (1987, pp. 646ff.) suggests that any properly principle-based (versus merely intuition-driven) deontology would not build in "exception" clauses allowing just the same things to happen; and he points to precisely such a phrase in Donagan's (1977, p. 82) *Theory of Morality*.

²¹ B. Williams 1973a, pp. 97–100.

²² Goodin 1982b, pp. 8–12. See also Hare 1981, pp. 48–9.

²³ Barry [1979a] 1989a.

and over the broad range of applications are also rules that broadly conform to the deontologists' demands.

This point is as old as the original utilitarian fathers who, while denying received moral rules any ultimate authority, nonetheless conceded that they might have derivative force insofar as they (or something very much like them) are sanctioned by the utility principle. In our own day, Richard Brandt has plausibly argued that the rules of war that we have inherited from the fundamentally deontological "just war" tradition are all broadly in line with what rule-utilitarianism would recommend.⁴⁴

Note carefully what I am arguing here. It is not that public officials will never experience utilitarian temptations to violate people's rights. The standard example for showing that concerns the case in which the only way to prevent a race riot that would kill dozens is by hanging someone whom we know to be innocent of any crime. My point is not that public officials will never face such situations, nor is it that they do not experience utilitarian temptations to violate people's rights (hanging innocent people, etc.) in such situations. My point is instead that public officials cannot systematically violate people's rights, as a matter of *policy*, and expect that policy to continue yielding the same utility payoffs time and again. Take the case of punishing criminal offenders, for example. The criminal sanction deters crime only in so far as it is imposed on the guilty and only the guilty. Introducing any probability that the innocent will be punished along with the guilty narrows the expected utility gap between criminal and noncriminal conduct, and increases the temptation for everyone to commit a crime. Thus, if we were as a matter of policy to punish people whether or not they were guilty, just according to some utilitarian calculation of public convenience on a case-by-case basis, then the utilitarian advantages of punishing the occasional innocent person would quickly diminish, and probably soon vanish altogether.⁴⁵

The reason utilitarian policy-makers are precluded from violating the rights of the innocent, as a matter of policy, is that policies soon become public knowledge. If nothing else, they are easily inferred

⁴⁴ Bentham [1789] 1970, chap. 2. R. Brandt 1972. This is, of course, just a special case of Sidgwick's ([1874] 1907) reconciliation of utilitarianism and commonsense morality.

⁴⁵ For a proof, see Goodin 1976, pp. 89–90, and, similarly, Rawls 1955, pp. 4–14. Likewise with Donagan's (1968, pp. 194–6) worries with the way utilitarianism would seem to give the lazy and penurious a right to soak the rich and industrious. The expectation of the extreme tax levels then in view would constitute a rational disincentive for the industrious ever to get rich; and that is why every political economist from Pigou (1932, pt. 4) onward has recommended that any redistributions come in the form of one-off lump-sum (rather than ongoing) transfer payments.

from past practices. Once news of such a policy gets out, people revise their expectations in the light of it – in the case of criminal punishment, their expectations of being punished even if not guilty. There are major utilitarian payoffs to be had from sustaining certain sorts of expectations and from avoiding others. Settled policies of one sort or another are characteristically required to produce socially optimal effects in both directions.* That is one reason for utilitarian policy-makers to abide by settled policies, even when greater utility gains might be achieved in any given instance by deviating from them.

Another, more pragmatic reason derives from “the argument from necessity.” Policy-makers, by reason of the circumstances under which they have to make their choices and the mechanisms they have available to implement them, are of necessity precluded from making any very fine-grained calculations. At most, they might make some very broad distinctions between different classes of cases; but picking out one particular case for special treatment is usually just not feasible. Policy-makers treat all cases according to some general rules because, in practice, they have little choice but to do so.

In response to the challenge that utilitarianism asks too little of us, then, it can be said that – at least as regards public policy-makers – utilitarianism demands not only about as much but also virtually the same things as deontologists would require. If they are going to decide cases according to general rules, rather than on a case-by-case basis, then the rules that utilitarians would adopt are virtually identical to those that deontologists recommend. And public policy-makers will indeed decide matters according to rules rather than on a case-by-case basis, either because the utility costs of doing otherwise are too high or else because as a purely practical matter more fine-grained assessments are impossible to make or to act upon.

C. Utilitarianism is too impersonal

A third line of criticism protests the impersonality of the utilitarian injunction. What should be done, according to its dictates, is simply “maximize utility.” Who should do it is simply everyone and anyone who is able to do so. On the face of it, at least, the utilitarian formula leaves little room for considerations of agency and virtually none at all for notions of “my station and its duties.” Capacities might vary, of course, and utilitarians could well understand how duties might naturally vary with them. But at root, utilitarian injunctions are utterly neutral with respect to agents. Utilitarianism supposes that, ultimately, the same injunctions apply to all alike.

* Hume 1739, bk. 3, pt. 2, sec. 2. Hodgson 1967. Harsanyi 1977b.

Critics of utilitarianism find this disturbing. Some duties, they say, must surely be agent-relative, with their content fundamentally (and not just derivatively) depending upon the agent to whom the demands are being addressed. I have a special duty to feed my *own* child, over and above (and separate from) whatever general, agent-neutral duty I might have to feed any hungry child that I might happen across. And so on.²⁷

How seriously this challenge should be taken, even at the micro-level, is once again open to question. Suppose that it is true that certain socially desirable consequences can be obtained only by assigning particular responsibility for certain particular performances to certain particular people. Then, clearly, utilitarianism would commit us to instituting just such a scheme to assign those particular responsibilities to those particular people. The duty to set up such a scheme in the first place might be a classically utilitarian one, a perfectly general duty shared by all alike. Once such a scheme is in place, though, the responsibilities assigned to particular people under it look very much like the "agent-relative special duties" that personalists so cherish.

They are not exactly alike, of course. But in so far as they differ at all, they differ principally in the very abstract matter of the ultimate source of their moral authority. For personalists, such agent-relative duties are something like moral primitives; for utilitarians, their moral force is derivative instead from the broader utilitarian considerations that guided their creation and allocation in the first place. Still, if the challenge is simply to account somehow for an obvious, first-order fact of moral phenomenology – namely, that each of us does indeed feel himself under different, special moral duties, over and above our more general duties of a baldly utilitarian form – then utilitarianism can indeed provide some such account.²⁸

In any case, the "impersonality" charge loses much of its force in the shift from micro- to macro-level applications of utilitarianism. Notice how many of the examples that antiutilitarians give of special, agent-relative duties are of a highly personalized sort. Duties arising out of kinship and personal commitments, for example, loom large. Almost all of the standard examples of special, agent-relative duties operate at this macro-level of personal conduct – and apparently necessarily so.

None of those standard examples of agent-relative duties survives transposition to the macro-level at all well. It may be wrong for a politician or civil servant, in his personal capacity, not to feed his own children; but it would be wrong for him, in his official capacity, to

²⁷ Nagel 1986, chap. 9. Sen 1982b. Parfit 1984, pp. 95, 485.

²⁸ Goodin 1985c. Pettit and Goodin 1986.

feed his own children before anyone else's in programs of disaster relief. It may be wrong for him, in his personal capacity, not to honor his promises or repay his debts; but it would be wrong for him, in his official capacity, to peddle influence, awarding public contracts to his past benefactors rather than to the lowest bidder. The classic sorts of special, agent-relative duties thus seem irrelevant at best (and utterly inappropriate, at worst) for guidance of officials in making and implementing public policies. Impersonality is not here a criticism. It is precisely what we expect of public officials in the discharge of their official duties.

In other ways, though, there are some respects in which personal commitments might properly guide policy-making at a more macro-level. One concerns the personal commitments of politicians – campaign promises and such like. Those arguably do give rise to special agent-relative duties that should be honored, even – indeed, especially – at the macro-level of public policy-making. A second sort of personal connection, analogous to kinship at the personal level, might be nationality at the macro-level. Just as parents are said to owe special, agent-relative duties to their own children, so too are we all – and our political leaders, as our representatives – thought to owe special, agent-relative duties to our compatriots.²⁴

The precise content of these supposed agent-relative duties, whether at the micro- or macro-level, has always been somewhat unclear. Depending upon what precise content they are given, advocates of agent-relative duties might find their case slipping away.²⁵ The particular distinction I would point to is between a duty “to do X” or “not to do Y,” on the one hand, and a duty “to see to it that X is done” or “to see to it that Y is not done,” on the other. As a shorthand, let us call the former “first-order duties” and the latter “second-order duties.”²⁶

The examples usually given of agent-relative duties are of the first-order kind. It is said to be an agent's duty to feed his own children, or not himself to kill or maim people, or some such. Indeed, the standard examples are emphatically *not* of the second-order sort. A parent who saw to it that his child was well fed by giving it up for adoption would be thought to have failed (or, anyway, repudiated) his agent-relative responsibilities in the matter altogether. Or, again, take someone who, like Jim in Bernard Williams' fanciful tale, saw to it that

²⁴ Compatriots are included in all the standard lists (see e.g. Parfit 1984, pp. 95, 185). Whether they should be is perhaps another matter (Beitz 1979, p. 163; Shue 1980, p. 132). Duties toward compatriots might just be an “assigned general responsibility” of the sort discussed above, as I argue in Chapter 17.

²⁵ Cf. Scheffler 1982; J. Bennett 1989.

²⁶ Or, more precisely, “second-order responsibilities” in terms of the distinction introduced in Chapter 2.

many people are not killed by himself killing one person. He would not ordinarily be said to have discharged his agent-relative duties on balance; indeed, he would not ordinarily be said to have discharged them at all.

Still, those are examples of agent relativity at the micro-level of personal conduct. Shifting to duties at the macro-level of public policy, agent relativity suddenly becomes much more of the second-order kind. No one seriously expects political leaders – even fabulously rich ones, like Kennedy, Rockefeller or Hoover – to feed starving compatriots out of their own larders. They are not expected to do it themselves, but merely to see to it that it be done. Neither, come to that, are we as a people necessarily expected to discharge our agent-relative duties toward compatriots through our own first-order actions. If the Dutch can manage to con the Americans into paying the whole cost of Holland's flood defenses, as some sort of NATO overhead, then the Dutch can hardly be said to have failed their agent-relative duties with respect to one another's physical security; rather, they have discharged them splendidly well. Thus, it seems that agent-relative duties, at the macro-level at least, require people to see to it that something be done or not done, not (necessarily) to do or refrain from doing it themselves.

Yet if the duty is merely to see to it that certain socially desirable consequences come about, then that duty fits perfectly well with a utilitarian ethic. What made special, agent-relative duties of the first order incompatible with utilitarianism was the insistence that you do or not do something, *regardless* of the larger consequences. The problem was that, in minimizing the evil done by your own hand, you may (if in the position of Jim or George) maximize the harm that comes about through others' hands. What makes special, agent-relative duties of the second order compatible with utilitarianism is their insistence that you do or not do something *because* of the consequences. Responsibilities to see to it that good outcomes obtain or bad ones are avoided – the sort of second-order special responsibilities that public officials ordinarily bear – are not counterexamples to utilitarianism. Instead, they are instances of it.

Having come this far, the only force now left to the impersonality criticism would derive from the fact that public officials are assigned special responsibilities for such matters, whereas utilitarians should presumably say that those are responsibilities which everyone should share. But utilitarians would insist only that everyone share them in the *first* instance, and in the *last*. In the first instance, everyone shares a utilitarian responsibility to maximize good. But if it turns out that more good could be done by assigning special responsibility to some particular agents for some particular matters, then that is what utili-

tarianism recommends be done. In the last instance, everyone is responsible for the consequences of sharing out responsibility in these ways; and if that moral division of labor has not worked successfully, then everyone shares the blame. In an ongoing social system, however, utilitarians usually will have rightly decided to appoint particular people to positions of special responsibility for certain morally important tasks. Public officials are a case in point. In ongoing systems, they will indeed have been assigned special, agent-relative responsibilities for seeing to it that certain morally desirable outcomes obtain or that certain morally undesirable outcomes be avoided, in a way that can be perfectly well rationalized on basic utilitarian principles.

III. REFUTATIONS REFUTED

Before closing, let me comment briefly upon two foreseeable lines of attack against the propositions that I have here been urging. My argument has been that public officials, in so far as they are utilitarians at all, are obliged to employ a form of utilitarianism that is less direct and hence more credible than the form that private actors would, by utilitarian principles, be obliged to adopt. Against the basic argument, I can anticipate two broad classes of objections.

One is that my argument amounts to saying that public officials are forced to adopt a form of utilitarianism that is inferior, by utilitarian standards, to that adopted by private individuals. Given that the critique would continue, it is surely better in purely utilitarian terms to leave as much as possible either to the actions of private individuals themselves or to their interactions through decentralized (e.g., market-like) coordination mechanisms. Surely it is better to put as little as possible into the hands of public officials, who are perforce obliged to do a worse utilitarian job in handling it.

The key question here is, of course, whether any less centralized system really would work any better to maximize social utility. There are some reasons for thinking that some might. Insofar as the problem with centralized systems is lack of information – insofar as that is why they are forced to work through general rules that fit poorly particular cases – there is the familiar argument, owing to Hayek, that markets are capable of taking account of much more of the relevant information.¹⁴

Set off against that, though, are all the familiar reasons that we have for supposing that we cannot accomplish through markets all that we achieve, however imperfectly, through more centralized systems.

¹⁴ Hayek 1945. Dasgupta 1982. pp. 205ff.

There are externalities that markets do not internalize; there are familiar problems of collective action associated with public goods, prime among them public order itself; and, perhaps most importantly in the present context, there are scale effects. If the value of an act repeated n times is greater than n times the value of the act performed once, then we have a particularly powerful case against decentralized (and in favor of coordinated, centralized) efforts to evoke such acts."

I am left concluding that we do, indeed, need centralized coordinating agencies issuing orders if we are to maximize social utility, and given the special circumstances surrounding them those agencies both ought and must issue orders that are general in form.

A second broad line of criticism would be that, while the form of utilitarianism adopted by public officials might usually evade criticisms rightly lodged against the form used by private actors, at least occasionally it will not. Perhaps public officials must usually act according to general rules, in order to induce and reinforce expectations that it is desirable from a utilitarian point of view that people should have. Suppose, however, that public officials can occasionally make an exception to those general rules without that becoming public knowledge. Then they can score act-utilitarian gains, while at the same time avoiding rule-utilitarian costs; so on utilitarian grounds it would seem better for them to deviate from the rules as necessary in such instances. Yet their deciding matters, even occasionally, on a case-by-case basis opens utilitarian public officials up to the same criticisms that are lodged against those who use utilitarianism as a guide to personal conduct.

Even if we can keep any particular secret, though, we will have a much harder time keeping secret our practice of keeping secrets.¹⁴ Should it ever become public knowledge that we break rules in secret (and hence unpredictable) ways, that would be even more deeply subversive of people's expectations and the utility predicated upon them than would be the knowledge of specific, patterned rule-breaking. The risk of those truly massive utility losses, notice, comes from the very act of keeping secrets at all. Thus, there is a very heavy surcharge, in utility terms, to be assessed against the very first secret deviation from general rules. And something of the same surcharge is paid with each successive secret deviation, if the risks of discovery of the general practice of secret deviation increase with the number of instances that there are awaiting discovery.

None of this constitutes an ironclad guarantee that there can never be a case in which the utility gain from deviating from a general rule

¹⁴ Harrod 1936, p. 148. Cf. Lyons's (1965) talk of "threshold effects."

¹⁵ Harrod 1936, p. 153. Goodin 1980, pp. 46-52.

is so great, and the probability of its being discovered so small, that public officials ought on utilitarian grounds secretly to break the rule. But this argument clearly does suggest that there is a substantial utilitarian presumption against such an action, and that the gains in view would therefore have to be very substantial indeed. Furthermore, where a truly exceptional case like this is involved, perhaps the utilitarian answer is the right answer after all." Here, as before, intuitions framed around more ordinary cases are likely simply to mislead.

IV. CONCLUSION

There has been no attempt here to try to anticipate every possible objection to utilitarianism, or to show how "government house utilitarianism" can be defended against each of them one by one. By canvassing the three major lines of contemporary criticism of utilitarianism, and showing how government house utilitarianism can be defended against all of them, though, I hope to have gone some way toward demonstrating the plausibility of that larger claim.

The basic trick, to be reiterated in all such defenses, is to draw a distinction between utilitarianism as a guide to personal conduct and utilitarianism as a guide to public policy-making, and to show that criticisms that are strong as applied to the former are weak as applied to the latter. What makes that claim plausible, in general, is the fact that public officials (both ought, and in any case must) govern through rules that are general in form.

Public policy-making takes place under some very special circumstances and operates through some very special instruments. Those special conditions pose special opportunities and special hazards. They also impose special constraints, not only on the forms of utilitarianism that public policy-makers can adopt but also on what alternatives to that utilitarianism they can reasonably be expected to contemplate. Under those special conditions that characterize public policy-making, utilitarianism looks distinctly credible, in a way it might not for private individuals in guiding their personal conduct.

" Even in utilitarian terms, though, it might be best to announce openly that the exception is being made - either incorporating that exception into subsequent formulations of the rule itself or, better yet (in light of the preceding arguments about the "peculiarities of public policy-making"), simply explaining why these were truly exceptional cases that should never be expected to recur.

PART III

Shaping private conduct

Responsibilities

Responsibilities are to consequentialistic, utilitarian ethics what duties are to deontological ones. Duties dictate actions. What responsibilities of the "task responsibility" sort advocated by utilitarians and consequentialists more generally dictate are, instead, results.¹ Exploring the deeper similarities and differences between the two notions helps us see what is at stake, and what is not (necessarily) at stake, in the larger choice between the two different styles of ethic.²

Of course deontologists do ascribe responsibility, in another sense, to agents. The sort of responsibility they are talking about is moral responsibility more generally – credit and blame for what one has done. I call this "blame-responsibility" in the chapters that follow. Utilitarian consequentialists need to ascribe to agents such credit and blame, too, if the task responsibilities they ascribe are to have any motive force. But the consequentialistic-utilitarian approach, by putting task-responsibility first and deriving moral responsibility from that, makes better sense of the way in which we actually assign credit and blame than does the deontological model, fixated as it is on motives and intentions and duty done for duty's sake.

This proposition is illustrated in the pair of chapters that follow. Chapter 6 considers the problem of ascribing credit and blame among participants whose joint contributions either overdetermine or underdetermine outcomes. Chapter 7 considers the problem of apportioning relative responsibilities for good or bad outcomes to participants in joint endeavors. In both cases, the account provided by deontological

¹ On different senses of "responsibility," see Baier 1970.

² The contrast points to theoretical options available to each kind of moralist that are denied to the other. In practice the contrast may be either stronger or weaker. Some consequentialists might choose to admit duties of a derivative sort; other consequentialists might choose not even to take up the option of ascribing responsibilities.

models seems decidedly inferior to that provided by models of task-responsibility.

I

Both duties and responsibilities are prescriptions of the general form:

A ought to see to it that X,

where A is some agent and X some state of affairs. In the case of duties, X takes the form:

A does or refrains from doing ϕ ,

where ϕ is some specific action. In the case of responsibilities, there is no reference in the X clause to specific actions on the part of A.

In the X clause describing responsibilities, there may well be a reference to the actions of others. Often our responsibilities require us to see to it that certain actions are performed or not performed. It is the essence of a military commander's responsibility to see to it that his troops attack enemy fortifications as instructed; it is the essence of the policeman's responsibility to see to it that people refrain from the various actions that constitute criminal conduct; and so on. But insofar as these are A's responsibilities, as distinct from A's duties, these are injunctions to see to it that *others* act or refrain from acting in certain ways. Call these "supervisory responsibilities." Such responsibilities take the form "A ought to see to it that X," where X denotes the state of affairs in which B does or refrains from doing ϕ and B is some agent not identical to A.

Or, again, the X clause describing A's responsibilities may well contain some essential references to A, just so long as it does not specify any particular actions which A must perform or refrain from performing. Thus, advocates of self-improvement may say that A is responsible for seeing to it that a state of affairs obtains in which "A's talents are fully developed." That counts as a genuine responsibility rather than a duty, provided that injunction is understood merely to set A a goal and leave open the choice of actions to be taken pursuant to that goal.

More typically, however, the state of affairs for which A is responsible will not refer to actions at all, or indeed to agents at all. A might be responsible for seeing to it that "the dog is fed," for example. The passive voice is significant, emphasizing as it does that what matters is the outcome and not the activity producing it. Perhaps A will feed the dog himself. Perhaps B will do it. Perhaps the dog will find its own food in neighbors' garbage cans or farmers' fields. No matter. So

long as Fido is fed, the state of affairs for which A is responsible will be said to have obtained.

11

Descriptions of A's duties must necessarily contain in their X clauses references to A's actions. Descriptions of A's responsibilities must not refer to *both* A and his actions, and need not refer to either, in their X clauses. They do, however, contain an essential reference to A and his activities just outside that X clause.

Responsibilities, by their very nature, require certain activities of a self-supervisory nature from A. The standard form of responsibility is that A *see to it* that X. It is not enough that X occurs. A must also have "seen to it" that X occurs. "Seeing to it that X" requires, minimally: that A satisfy himself that there is some process (mechanism or activity) at work whereby X will be brought about; that A check, from time to time, to make sure that that process is still at work, and is performing as expected; and that A take steps as necessary to alter or replace processes that no longer seem likely to bring about X.

These self-supervisory responsibilities are marked off by a special vocabulary used to characterize their breach. When an agent A has failed to produce the state of affairs X, we say merely that "A has failed to discharge his responsibilities." But when A has failed to discharge these self-supervisory responsibilities we accuse him of the much graver offense of being "irresponsible," that is, insensitive to his responsibilities. And that charge would stick even if, through some fortuitous circumstance, X came about even though A took no steps to see to it that it did.

Nor, positively, could A be said to have discharged his responsibilities in the fullest sense unless he "saw to it" that X obtained. In the case of A's responsibility for seeing to it that Fido is fed, A has not necessarily discharged his responsibilities whenever Fido has a full belly. His responsibility may be irrelevant – devoid of action-implications – under those circumstances. But A could hardly claim to have *discharged* his responsibilities just on the grounds that X obtains, if he had never even stopped to check whether or not Fido was hungry.

These self-supervisory responsibilities are genuine responsibilities. Although directed at A himself, they are injunctions that mandate goals and very general classes of activities, rather than specific actions. There are various ways A might set about satisfying himself that some process is at work to bring about X, or checking that it is still working properly, or changing it if it is not. It is by virtue of this flexibility in the choice of specific actions to be performed that these self-

supervisory requirements can properly be called "responsibilities" rather than "duties."

Notice, furthermore, that this self-supervisory requirement is a feature which responsibilities share with duties. Those, too, are characterized by the general formula "A ought to *see to it* that X," where X is understood as "A does or refrains from doing ϕ ." What A is required to exercise self-supervisory responsibilities over differs in the case of duties. There, instead of outcomes, the focus is on A's own actions and inactions. But clearly it is not enough, for A's duty to have been discharged, that the right action was performed or omitted. To do his duty, A must furthermore *see to it* that this is the case. That is to say, A must ϕ or refrain from ϕ ing consciously, intentionally and purposively in order to truly have done his duty. Doing what is required accidentally, or incidentally to the pursuit of some other goal, or as a result of posthypnotic suggestion, does not qualify as doing one's duty in the fullest sense.

Even in the case of duties, however, this self-supervisory requirement itself takes the form of a responsibility. All that this requirement demands is that A *see to it* that A ϕ s or refrains from ϕ ing. While the duty itself – the X clause – demands specific performance of some particular kind, the self-supervisory requirement does not. There are many ways A might go about "seeing to it" that he does his duty. The choice between them is left to his discretion. What the self-supervisory "see to it" requirement mandates is a result – that A do or refrain from doing ϕ – rather than any particular actions that A must take to achieve it. That being so, the self-supervisory requirement built into duty statements itself qualifies as a responsibility rather than a duty.

III

Like duties, responsibilities constrain the bearer. In the former case, A is duty-bound to perform (or refrain from performing) certain actions. In the latter, A is bound by his responsibilities to strive for certain goals. What crucially differentiates responsibilities from duties is the discretionary component necessarily built into them. Whereas duties require specific performance of certain actions, responsibilities allow agents to choose between alternative actions having the same overall consequences. A is responsibility-bound to *see to it* that X. But he is not told how to go about doing that. It is left to his discretion.¹

In certain circumstances, it may seem that fixing A's goal determines which course of action he is morally obliged to pursue. Suppose

¹ See similarly Feinberg 1966, p. 141.

A has responsibility for seeing to it that some state of affairs X obtains, and the only way (or perhaps just the only way under his control) for that result to be produced is for him to ϕ . Then A's responsibility to see to it that X would seem to translate into a duty that he see to it that he ϕ .

Notice, however, that any duty thus derived is contingent in various ways that a deontologically dictated duty for A to ϕ would not be. If A is required to ϕ merely because that is the only way for him to see to it that X, then that is a mere happenstance – a quirk of the causal machinery as presently arranged – rather than a necessary feature of any and all possible worlds. Should some new way arise whereby A can see to it that X obtains, then the responsibility in view would no longer entail a duty for him to ϕ . A's ϕ ing would have become optional, in a way that it never could be were his duty to ϕ deontologically derived. Furthermore, were A's duty to ϕ derived from his responsibility to produce a certain state of affairs which is contingently connected to his ϕ ing, there would be nothing wrong (and perhaps much right) with him trying to find alternatives to ϕ ing. Were A's duty to ϕ deontologically derived, such behavior would constitute a clear attempt at evading the duty. Thus, even if ϕ ing were the only way A could see to it that X, the implications of ascribing a responsibility to see to it that X still differ importantly from those of ascribing a full-fledged duty to ϕ .

IV

Finally, the discharge of a duty is a binary variable, whereas the discharge of a responsibility is a scalar one. A duty is here analyzed as an injunction to perform some specific action. Either A has ϕ ed (in which case he has discharged his duty), or else A has not ϕ ed (in which case he has not discharged his duty). Assuming ϕ and its negation logically exhaust all the possibilities, there is simply no scope for saying that A could ever have "more-or-less" discharged his duty.

With responsibilities, there is considerably more scope for saying just that sort of thing. In the case of a responsibility, A is enjoined to see to it that certain results described in the X clause are produced. Of course, if A has seen to it that precisely those results obtain, then he will be said to have discharged his responsibility fully. But suppose now that A has seen to it that results are produced which, while not exactly the results specified in the X clause, are substantially similar to those results. Then he can be said to have discharged his responsibility more or less completely, depending on the extent of the similarity between the results produced and those mandated.

The reason responsibilities can be discharged more or less com-

pletely, and duties cannot, derives, in turn, from a deeper feature of the two different ethical systems in which those notions are set. Responsibilities admit of partial discharge because, in a consequence-based ethic, different outcomes can be more or less substitutable for one another. The discharge of a duty is, in contrast, an all-or-nothing affair because, in the deontological ethic from which such duties derive, there is no substitute for doing what your duty requires you to do. In failing to discharge one duty, you might of course discharge another duty that is almost as important. But what you will then have done is not describable as "almost" having discharged the former duty. You will instead be said to have failed to discharge it at all, although having fully discharged another duty which is almost as important.

Some responsibilities can, at least in principle, be discharged completely. Others by their very nature cannot be. A's responsibility to see to it that his dogs (or children) are fed, for example, is discharged fully and completely whenever the dogs (or children) are indeed fed, and A has seen to it that this is so. Let us call this a "fixed-target responsibility." The target – a dog or child that is "well fed," whether defined in terms of minimum daily requirements or satiety – is a fixed target. It is therefore possible, at least in principle, for A to meet (or indeed beat) that target and thereby discharge his responsibilities fully and completely.

In practice, of course, the target may be fixed so high and A's resources so low that he cannot meet the target. Then A will be said to have discharged his responsibilities with respect to the fixed target more or less completely, depending on the extent to which he has seen to it that his dogs or his children are fed as well as he is able.

The notion of a more or less complete discharging of responsibilities really comes into its own, however, in relation to another type of responsibility. Call these "receding-target responsibilities." There, A bears a responsibility:

A ought to see to it that X,

where X describes some ideal state of the world which can never be fully attained, but which can only be approximated. Since the mandated state of affairs can never be completely attained, A can never fully discharge his responsibility. Were responsibilities all-or-nothing propositions like duties, it would make no sense to enjoy the impossible. But since responsibilities admit of more or less complete fulfillment, it makes perfectly good sense to describe our responsibilities by reference to some impossible ideal, and then to ask not *whether* A has discharged his responsibilities but rather *to what extent* he has done so.

The similarities between duties and responsibilities bear as much emphasis as do the differences. What those similarities, taken together, suggest is that there is less than is commonly supposed to choose between in the consequentialist-deontologist debate. Important differences do remain, but not nearly as many as existed before we came to appreciate the way in which responsibilities can serve as the functional equivalents of duties.

Where deontologists would assign duties, consequentialists would – for perfectly good consequentialistic reasons – assign responsibilities. Not everywhere, perhaps, given their different grounds for constraining people's choices; but over the broad range of cases this would be true. Furthermore, these responsibilities can be just as universal or just as agent-specific as duties can be. And in their material consequences, the practical upshot of both brands of constraint is broadly similar: not identical, again, but substantially the same in a substantial number of cases.

This offers utilitarians valuable opportunities for taking on board some of the more important insights of their deontological critics without abandoning their consequentialism in any respect. At the same time, it provides them with a valuable way of defending themselves against certain sorts of deontological attack. The machinery of responsibilities is available to them to do some of the sorts of things that deontologists rightly demand that any plausible ethic should be able to do. Whether the machinery of responsibilities or the machinery of duties is better suited to any particular task is an open question. But it is an important and less contentious preliminary to that larger argument to point out that consequentialists do not come to those tasks empty handed.⁴

⁴ This chapter grows out of work on my book, *Protecting the Vulnerable* (Goodin 1985c), and I am grateful to many people who through comments on that manuscript helped indirectly to shape this one. I am particularly grateful to Philip Pettit, a collaborator on another paper (Pettit and Goodin 1986) which significantly shaped this chapter.

Distributing credit and blame

My larger aim in this chapter and the next is to demonstrate the relative inadequacy of deontological models of blame-responsibility compared to consequentialistic utilitarian models of task-responsibility. In opening that attack, let me first focus on the inadequacy of models of blame-responsibility where they might be expected to be on firmest ground.

Dictating how to distribute credit and blame among agents for what they have done or not done is, on the face of it, something that models of blame-responsibility should be able to do well. In truth, those models perform that task quite poorly, at least in an arguably common class of cases. Where the credit or blame is being distributed among people whose actions jointly either underdetermine or overdetermine the outcome, models of blame-responsibility are radically incomplete guides to the distribution of credit or blame. At most, they fix only broad parameters for those distributions.

My analysis on this point begins with some commonplace observations. Outcomes are characteristically the products of actions and omissions of many people. Furthermore, different people make different, and differentially important, contributions to the end product. Both these facts are well mapped in our standard notions of "responsibility." Whatever particular account we want to give of that notion, it is clear that responsibility can be shared, and shared unequally, between various different agents.

What is less familiar, but nonetheless true, is that these fractional responsibilities do not necessarily sum to one. Sometimes they sum to more than one. For a familiar example, consider how each of several coconspirators is held fully responsible for the crime that they have committed together. There is one dead body, say, but nine people fully responsible for the murder. Other times, fractional responsibilities sum to less than one. An example here might be the responsibilities

of a pair of reckless drivers, each of whose negligence contributed to (but did not, separately or together, fully determine) the wreck on the highway. There is one wreck, say, but two people, each of whom is one-quarter responsible for it.

The first task of this chapter is to trace out some of the circumstances in which, and reasons for which, fractional responsibilities fail in such ways to sum to one. The second and larger task of this chapter is, building on the observation that that does sometimes happen, to consider what its implications might be for theories attempting to derive people's deserved treatments from their responsibilities.

Moral theories based on deontological notions of desert or blame-responsibility standardly try to analyze how we should treat (reward or punish) people directly in terms of the moral merits of their past performances.¹ In the circumstances here in view, which are arguably common circumstances, that is strictly impossible. A further bridging principle is required. And the only one that is available, within deontological models of blame-responsibility, is peculiarly weak and (as will be shown in Section V) ultimately incoherent.

For the purposes of this chapter, I make no distinction between responsibilities for "positive" and "negative actions" – between actions and omissions, what one has done and what one has failed to do. Arguably the latter are always morally on a par with the former, anyway. But in any event, that seems undeniably true where, as I shall endeavor to ensure is clearly the case in all my examples here, the omission constitutes a failure to discharge some prior, independent moral responsibility. Likewise, my account will be symmetrical as between responsibilities for "positive" (good, desirable) and "negative" (bad, undesirable) outcomes.

I

The argument of this chapter hangs on the proposition that fractional responsibilities do not always sum to one. Their summing either to less or to more makes trouble – and trouble of a strictly analogous form – for the attempt to derive deserved treatments from statements of fractional blame-responsibilities. For the purposes of this preliminary demonstration, however, let us take the two cases separately.

¹ There are of course exceptions – notably, cases of vicarious liability where an employer is liable for any damage done by his employees or a dog-owner by any damage done by his beast without being himself in any way causally implicated in the harm (Hart 1968; Feinberg 1970). But those exceptions seem so exceptional as to confirm the converse as the general rule.

A

The clearest cases of fractional blame-responsibility summing to more than one are cases in which the outcome has been overdetermined. Suppose that A has been administered a slow-acting but utterly lethal poison by B, but just as B's poison begins to work C shoots A dead. Then we would not only want to hold B and C equally responsible for A's death. Surely we would want to hold each of them *fully* to blame.² That, however, would mean that B and C's fractional responsibilities for A's death sum to two, not one.

The crucial thing that makes us want to hold both parties fully responsible for an outcome that is overdetermined by their actions is just this. The action of each of them, taken independently of the action of the other, was *fully sufficient* to determine the outcome.

Similar conclusions might obtain with respect to responsibilities for omissions. Fractional responsibilities might sum to more than one where there are several agents omitting to act, any one of whose acting instead would have been sufficient to produce (or, more typically, avert) a certain outcome. Suppose A is a poor swimmer who is obviously drowning at the seashore equidistant from lifeguard B and lifeguard C. Either B or C (or both of them) could have waded in and rescued A. If neither does, we are inclined to say not only that they are equally responsible for A's drowning. We would surely want to hold each of them *fully* to blame for A's death.³ Again, B and C's fractional responsibilities sum to two, not one.

Once again, the reason for this judgment is that either party's action, taken independently of the other's, would have been sufficient to avert the outcome in view. Another way of putting that point is to say that the contribution of each was *fully necessary* for the outcome's occurring.

Both those are cases in which agents make their causal contributions independently of each other's. Consider next cases of interdependent actions.

Note first that fractional responsibilities may sum to more than one where an agent induces another to act. Suppose that A has been shot dead by B, who is a contract-killer who has been hired by C to kill

² The example is from Mellema (1984). It is well established, both in ethics and law, that a conspirator's responsibility is nowise diminished the more coconspirators there are (Pufendorf 1717, bk. 3, chap. 1, sec 5; Nozick 1972, p. 130; Zimmerman 1985, pp. 115-22; Goodin 1985c, pp. 135-6). Similarly in tort law, joint tort-feasors are held jointly and separately liable for the full cost of repairing any damage they have done together (G. Williams 1951, chap. 1; Prosser 1971, chap. 8).

³ Feinberg 1970, p. 244.

A.⁴ In such circumstances, there is surely a case for saying that B and C should each be held fully responsible for A's death; and if they are, then fractional responsibilities once again would sum to two rather than one. But even if we do not want to go quite so far as to hold the gunman and his employer equally responsible for the death, it is nonetheless clear that B and C's fractional responsibilities, taken together, must sum to more than one. This follows because we do not think that the responsibility of the hired gunman, B, is in any way diminished by the fact that he has been hired by C – B is *no less* responsible for A's murder than would be some assassin (call him B') who acted at his own initiative. But we also want to hold C, who hired the gunman, responsible in *some* measure for A's death. Hence, the fractional responsibility falling to B, summed with that falling to C, must exceed one (if that is properly reckoned as the quantum of responsibility that would have fallen to the unhired gunman B').

The reason here that responsibilities sum to more than one is a conjunction of the two types of reason offered in the analysis of the two cases described previously. One crucial factor is that people whose actions are *sufficient* to produce an outcome are fully responsible for that outcome, the shared responsibilities of others notwithstanding. A second crucial factor is that those whose actions are *necessary* to produce an outcome are not exonerated (not fully, anyway) by the fact that others also make, and must make, causal contributions – even if those contributions are later or in some sense 'larger' ones.

The analysis just offered was cast in terms of actions. But nothing in it changes if we recast it in terms of omissions instead. In the story as told, for example, C commissions gunman B to *act* so as to cause A's death. Let us vary that aspect of the story: suppose B is instead the sole lifeguard employed at A's personal pool; A is a miserable swimmer who requires constant rescuing; C, in these circumstances, commissions B merely to *omit* to rescue A the next time he is floundering. The conclusion that fractional responsibilities of B and C for A's death sum to more than one remains unchanged in these altered circumstances. B is no less responsible for A's death than would be a B' who omitted a similar duty without financial inducement; C bears some responsibility for A's death; hence B plus C's responsibilities exceed one (that being reckoned as the responsibility falling to B').

The previous examples are ones of interdependent actions involving complicity, of one form or another, between the two parties jointly

⁴ The same may be true of positive actions producing good results. Consider the case in which A is a tennis player trained by B, a coach hired by C to train A.

involved in producing a state of affairs for which they are both to be held *ex post* responsible. But that is nowise crucial. Fractional responsibilities can sum to more than one even where that feature is missing.

Consider first cases of "passing the buck." There, the agent passing the buck intentionally engages the other in the activity; but the one to whom the buck is passed did not himself intend to get involved. Suppose, for example, that an unscrupulous firm engages in a productive process that it knows will turn a quick profit but which it also knows will create long-term harms for its employees, customers or neighbors. Suppose that, knowing this, the firm takes its profits early and declares itself bankrupt at just the moment the costs are starting to accrue, confident that the government will then step in to clean up its mess for it. Suppose, however, that the government did not do so.³ Then we would surely want to say that both the firm and the government were responsible – the firm *fully* responsible, and the government at least partially so – for the harms that ensue. And, of course, one agent's full responsibility plus another's partial responsibility adds up to more than one.

Conversely, there are cases of "setting a bad example." There, the agent setting the bad example acts in the foreknowledge, but without the intention, that his act will influence that of another. Here, it is the action of the one who is following the bad example that is intentional, instead. Suppose the United States dumps nuclear wastes at sea, knowing but not intending that other nuclear powers will follow its lead; suppose, furthermore, that perfectly predictable ecological damage ensues as a result of all this foreseeable dumping that follows from the initial American act. In such a case, we would surely want to say that each nation is fully responsible for the damage that its own waste does. But, over and above that, we would want to say the United States is responsible for setting a bad example. In addition to being fully responsible for the effects of its own wastes, it is also partly responsible for the effects of the wastes that it unintentionally but knowingly encouraged others to deposit. Once again, full responsibility plus partial responsibility adds up to more than one.

The crucial thing making fractional responsibilities sum to more than one, in both these cases, seems to be this. Although you as the "innocent party" did not intend that the other should shape his behavior to your own, you can reasonably be expected to know that he

³ Or, for another example from an earlier era, suppose that parents default on the obligation they assumed at their child's christening to care for the spiritual well-being of the child, expecting the godparents to do so instead; but they do not, either. Then both the parents and godparents would be fully responsible for the child's spiritual plight.

the other has done or will do so. The other's intentional shaping of his action to yours lets him in for full responsibility for the outcome of your joint actions. Your knowledge that he has done or will do so lets you in for partial responsibility for the selfsame outcome. Together, those responsibilities sum to more than one.

B

Cases in which fractional blame-responsibilities sum to less than one are relatively straightforward, in contrast. They consist essentially of cases in which outcomes are underdetermined by the joint actions of the several agents involved. That is to say, the actions of the parties were neither individually necessary nor jointly sufficient to produce the outcome; the intervention of some further agent, generically characterized as Chance, was required to produce the outcome that eventuated.

Consider the standard case of an automobile crash. The crucial factors in contributing to the crash, let us say, were driver A's slightly bald (but perfectly legal) tires, driver B's slightly weak (but perfectly legal) eyesight, and an unusually heavy fog. Certainly individuals A and B's contributions to the crash were not themselves sufficient, individually or jointly, to produce the crash; had there been no fog, let us say, there would quite assuredly have been no crash. Let us imagine (as often presumably would be the actual case) that neither was necessary, either. That is to say, the fog was so bad that there was a fair chance that the two drivers would have hit one another's car, even absent problems with tires or eyesight; all those factors did was increase an already nonnegligible probability of an accident.

Under such circumstances, it seems distinctly odd to suppose that the fractional responsibilities of the two drivers sum to one. We may quite sensibly say that B is somewhat *more* responsible than A, since B should have known his eyes were weak (e.g., he has to squint when reading signs in the distance) whereas A had no particular reason to suppose his tire tread was wearing thin (although of course he should have checked it more often, just as B should have had his eyesight checked more often). But all that is beside the present point. What is at issue here is whether the fractional responsibilities of A and B, however apportioned between them, sum to one.

Where parties' faults, taken together, come so far short as they do in this instance of determining the outcome, I think it unreasonable to say that fractional responsibilities sum to one. The only time that would make sense, I suggest, is when a pair of agents, given their combined faults, could be said to deserve to have some misfortune

befall them – when they can be said to “have it coming.”⁶ Whenever we can properly say that the pair of them, though slightly negligent in certain respects, were nonetheless simply *unlucky* that the untoward event occurred, I suggest that a large portion of the responsibility should be assigned to Chance. The individuals, taken separately, should be allocated responsibility for such outcomes as would characteristically follow from their faulty performances. If, through some fluke, something much worse happens, then it is not properly chargeable to their account.

II

Thus we have seen that, in many diverse (and arguably common) circumstances, fractional blame-responsibilities do not necessarily sum to one. Sometimes they sum to more, sometimes to less. In either case, the upshot is the same. Any direct link between responsibility and liability (i.e., deserved treatments, be they rewards or punishments) is necessarily severed.

The reason this is so is straightforward. Consider first a case where fractional responsibilities sum to more than one. Imagine ten individuals, each of whom is to be held fully responsible for causing damage to a jetty that will cost \$100 to repair. Were we to derive liability directly from blame-responsibilities, saying that each of the ten was fully responsible for the \$100 damage would entail collecting \$100 from each of the ten. You would then have collected a total of \$1,000 to effect repairs costing \$100. Who deserves to get the extra \$900? The injured party has done nothing to deserve it. Neither has anyone else.

Or, again, suppose that each of ten people makes an absolutely essential contribution toward the production of a widget valued at \$100. Then, too, we would say not just that each of the ten was equally responsible but, following the discussion in Section IA, that each of them was fully responsible. But were deserved treatments derived directly from responsibilities, saying that each of the ten was fully responsible for the \$100 widget would entail paying each of them \$100. You would have ended up paying \$1,000 in remuneration for the production of something that can be sold for only \$100. Where is the extra \$900 to come from?

Similarly in cases of fractional responsibilities summing to less than one, it makes no sense to derive liabilities from responsibilities directly. If there are ten hikers, each of whom is 1/100th responsible for harm to an embankment costing \$100 to repair, then equating liability with responsibilities would lead us to bill each for \$1. You end up

⁶ Goodin 1988, chap. 10.

collecting a mere \$10 that way, however; \$90 is still outstanding. Presumably that bill is chargeable to Chance. But where do you address bills that are to be sent to Chance? Or, again, suppose that ten people working together discovered some fantastic new drug, but that they did so mostly through sheer good luck. Each of the ten can only be said to be $1/100$ responsible for the discovery, let us imagine; the rest of the credit goes to Dame Fortune. Then equating deserved rewards with responsibilities would lead us to pay each of the ten 1 percent of the profits from sale of the drug. The other 90 percent is surplus. Who deserves to get it? *Ex hypothesi*, all those who have done anything to deserve it have already been paid in full.

III

The most that can be said would seem to be this. Where fractional blame-responsibilities fail to sum to exactly one, a person's responsibility will set an upper bound (when they sum to more than one) or a lower bound (when they sum to less than one) on what treatment that person should receive. By "upper bound" I mean to say that people ought not be assessed penalties or rewards that are in excess of that sum. By "lower bound" I mean to say that people ought not be assessed penalties or rewards that are less than that sum.

Consider the case of the ten yachtsmen who are each fully responsible for causing \$100 damage to a jetty. On this analysis, each may properly be held liable for up to \$100 in damage payments, but no more. That upper bound leaves open, of course, a broad range of permissible alternative cost-sharing arrangements among the yachtsmen. These range from charging any one of them the full \$100 and all the rest nothing right the way through charging all of them \$10 equally. None of these schemes violates the upper bound, so any of them would be permissible. Given that the yachtsmen's fractional responsibilities sum to more than one, that is the most that analysis of their fractional responsibilities alone will tell us. There is no other way of deciding, in those terms alone, which of these many possible cost-sharing schemes is to be preferred.

Similarly in the case of the widget-makers, each of whom is fully responsible for causing a \$100 widget to exist. Since their fractional responsibilities sum to more than one, there is an upper bound of \$100 upon what can properly be paid in remuneration to any widget-maker. But, again, there are a variety of possible alternative remuneration schemes, ranging from paying one the full \$100 and the rest nothing, right the way through paying all of them \$10 equally. None violates the upper bound; all would be permissible. Given that the widget-makers' fractional responsibilities sum to more than one, that

is the most that any analysis of their fractional responsibilities alone will tell us.

In cases where fractional responsibilities sum to less than one, there is a lower bound set instead of an upper one. Each of the hikers should be liable for at least \$1 in damage payments for the harm they have done. Each of the pharmacologists should properly receive at least 1 percent of the profits from the drug. We may decide to charge (or pay) them much more than that, if we happen to believe that costs and benefits of accidental events should be borne by those most directly involved. But that amounts to assessing liabilities *in excess of* – not in line with – responsibilities. Alternatively, we might suppose that society should pay for accidental damage out of general fund revenues (and deposit the gains from accidental benefits to general fund revenue). Either of those options – or anything in between them – would be perfectly permissible. Given that fractional responsibilities here sum to less than one, the most that an analysis of people's fractional responsibilities alone can tell us is that we must respect the lower limit.

IV

When fractional blame-responsibilities fail to sum to one, then, we face a broad range of permissible cost- and benefit-sharing options. How we choose between them is in no way dictated by an analysis of people's fractional blame-responsibilities alone. We rely instead upon other blatantly utilitarian principles. We may, for example, allocate liability for damages to whomsoever has the "deepest pockets," that is, whomsoever can best afford to bear the costs. In so doing, we do not pretend to apportion liabilities proportional to moral blame, responsibility or anything of the sort. We are not *blaming* anyone for having deep pockets, nor are we accusing him of any *wrongdoing* in the process of acquiring deep pockets. But we do not have to do so, either. For we have seen that the analysis of responsibility is silent as between alternative allocations of liability within the broad ranges picked out.

Neither do we pretend that, once all the \$100 damages have been compensated to the owner of the damaged jetty, that those yachtsmen who escaped paying but who were fully responsible nonetheless are off the hook. They are not liable for anything, perhaps. The jetty-owner, having been fully compensated, has no further claim. But the culpable yachtsmen remain responsible, even if their liability is at an end.

Similarly in the case of allocating rewards, fractional responsibilities summing to something other than one open a broad range of permissible cost- and benefit-sharing options. You need not try to pay

people in strict proportion to their fractional responsibility, where many share full responsibility for producing the product (as, e.g., capital and labor typically do, if you believe in Cobb-Douglas production functions). Rather, you are freed to largely ignore questions of desert and responsibility and distribute benefits on other bases entirely – need, for example – just so long as you do not pay anyone more than the upper limit suggests. Here again, you do not need to pretend that widget-makers who were responsible but not rewarded their full \$100 have had their claims canceled, any more than those yachtsmen who were negatively responsible for damaging the jetty but who escaped the bill have had their liabilities canceled. Their responsibility for producing the widget remains unquestioned. It is just the liability – the reward, here – that has varied.

V

Of course, even where fractional responsibilities do not sum to one, we might nevertheless use them to allocate liability (reward, punishment) a little less directly. There can be no question of setting fractional allocations of liability *equal* to fractional responsibilities. Since fractional responsibilities sum to something other than one, we would by doing that end up allocating either more or less than the total sum of liability to be distributed.

Still, if we want to allocate total liability according to fractional responsibilities, and those alone, we can always set fractional allocations of liability *proportionate* to fractional responsibilities. Suppose individual A bears $1/a$ responsibility for the outcome and individual B bears $1/b$. The proportionality rule suggests that the ratio of A's liability to B's (L_A/L_B) should be the same as the ratio of A's responsibility to B's ($\frac{1/a}{1/b} = b/a$), and that this ought to be true even if $1/a + 1/b \neq 1$. This proportion is then used to allocate total liability. So, for example, if two people are each fully responsible for the death of some third party (as are the poisoner and the shooter in my initial example), they would each bear liability for half a murder. And if there were a third party who was also fully responsible for the murder, they would each bear liability for a third of a murder. And so on.

The appeal of this proportionality rule presumably lies in this "fairness," "equity" or "justice." If two people have worked identically to produce the same outcome, then they should in all justice pay the same penalty or reap the same reward.

The notion of justice to which the proportionality rule appeals is,

obviously, what Feinberg would call a "comparative" standard of justice.⁷ But there is another, noncomparative standard of justice that could also be applied to such cases. Certainly it is not unjust to judge someone fully liable for the penalties we inflict on murders if that person is fully responsible for committing a murder. The fact that others are also fully responsible for the same murder, but escape liability for it, in no way impugns the (noncomparative) justice of holding the one party liable for it, and fully so. That no more follows than does the conclusion that just because we cannot catch every criminal we ought not punish any. It may be unjust to do so according to standards of comparative justice. But it certainly is not unjust according to the more appropriate standards of noncomparative justice that should be used in such cases.

It would, of course, be unjust even in noncomparative terms to move outside the bounds set by people's fractional responsibilities themselves. If there are three people, each of whom bears half responsibility for causing certain damage, then it would be unjust even in noncomparative terms to charge any of them for more than half the repairs. But there would, in noncomparative terms, be nothing unjust in charging two of them each one-half of the costs, and the third nothing.

Not only does the proportionality rule go beyond what noncomparative justice rightly demands. It also offends against its own preferred standard of comparative justice. Its basic premise is that like cases should be treated alike, and different cases differently in proportion to their differences. For those recommending the proportionality rule on these grounds, the focus is on the liabilities that fall to two people who are identically involved in the same activity together. It seems (comparatively) unjust that one should bear more liability than the other in such cases.

But let us shift the focus slightly. Consider instead the liabilities of two people. One of them, let us imagine, is fully responsible for a murder because he poisoned someone who then died of the poison. The other is fully responsible for murder because he poisoned someone who would have died from that poison if someone else had not shot him dead first. These two murderers have performed identical acts with identical intentions, let us imagine; so their liabilities, on the standard of comparative justice here in view, really ought to be identical. Yet the proportionality rule would allocate the second poisoner half as much liability as the first, on the grounds that his liability must be split equally with the shooter. Thus, the proportionality rule produces comparative justice between those responsible for the *same* mur-

⁷ Feinberg 1974: 1980, chap. 13.

der at the cost of comparative injustice between those identically responsible for *different* murders.

The upshot would seem to be that comparative justice is an incoherent standard, in such cases. However we allocate liability we would offend against it, in one way or another. But comparative justice is the only principle to which the proportionality rule can appeal. If it is inapplicable (because incoherent) in cases where fractional responsibilities sum to something other than one, then we must dismiss the proportionality rule along with it.

VI

The arguments of this chapter trace an important limit to the model of blame-responsibility, even on its home ground. These arguments seem to suggest that questions of reward and punishment, praise and blame, really are largely separate from questions of blame-responsibility, at least for the arguably broad class of cases where fractional responsibilities do not add up to one. Not only are deontological models of blame-responsibility incomplete, in such cases. Furthermore, the only way they can be made complete – through appeal to the proportionality rule of comparative justice, discussed in Section V – is ultimately incoherent.

Thus, even in the core area of their application, deontological models of blame-responsibility not only may but must be supplemented by more consequentialistic-utilitarian models of responsibility. Being forward- rather than backward-looking, aiming at securing desirable outcomes for the future rather than doing justice to people's past performances, such consequentialistic-utilitarian models of responsibility are largely immune to the forces that make deontological models of blame-responsibility so radically incomplete.

Apportioning responsibilities

The preceding chapter was essentially a softening-up exercise, showing that the deontologists' model of blame-responsibility was incomplete even in its own terms. Next, I want to mount a more positive case for the relative superiority of consequentialistic utilitarian accounts of task-responsibility over deontological models of blame-responsibility, based on the problem of apportioning responsibilities among participants in some joint endeavor.

The argument of this chapter starts from many of the same commonplaces as that of the preceding one. Maybe occasionally individuals act alone, one individual bearing exclusive responsibility for some particular state of affairs. More commonly several individuals act together, "sharing responsibility" (in some sense or another) and, indeed, sharing it differentially. There, responsibility is neither exclusive nor equal. Some people bear rather more responsibility, others rather less, for one and the same state of affairs.

That is simply to say that we characteristically find ourselves *apportioning* responsibility between different agents. In itself, that observation seems unexceptional and unexceptionable. Upon further inspection, however, it turns out that that observation is not nearly so idle as it seems. That we apportion responsibility, and how we do so, in the end tells us something important about the nature of responsibility itself. For it is only on certain models of moral responsibility – the standard model, extraordinary as it seems, not being among them – that that practice makes sense.

The standard model of responsibility (championed variously by lawyers, churchmen and Kantian philosophers) is essentially a mechanism for fixing credit or, more commonly, blame for certain sorts of states of affairs. Hence I shall dub this model one of "blame-responsibility." People are paradigmatically held responsible for flaws in their characters and for the outcomes (especially the harms) that

flow from them. People who harbor evil intentions, producing harms on purpose and by design, are charged with the gravest form of responsibility, marked off by the criminal law and its moral equivalents; those who are merely negligent, unintentionally but recklessly causing harm, are charged with lighter forms of responsibility, marked off by the civil law and its moral analogues. On both levels, the central concern of advocates of the model of blame-responsibility is with "what sort of person one is," as revealed through one's actions and what (motives, thoughtlessness, etc.) can be seen lying behind them.

Presumably ascribing responsibility is always essentially a matter of pointing fingers. What distinguishes the standard model of moral responsibility from the one that I shall here be advocating is the *purpose* for which fingers are being pointed. With the standard model of responsibility, fingers are being pointed for purposes of fixing blame. With my model of responsibility, fingers are being pointed for purposes of assigning duties, jobs or (generically) tasks. Hence mine is a model of "task-responsibility." In it, questions of "who has been assigned what tasks?" become truly central to the business of ascribing responsibilities. The difference between "my job" and "your job," on this account of responsibilities, looms particularly large.

Of course, it would make no sense to assign people task-responsibilities if they were not also going to be held to account for how well or badly they performed those tasks and discharged those responsibilities. Therefore, task-responsibilities inevitably entail correlative blame-responsibilities. But the blame-responsibilities entailed by task-responsibilities are different, both in substance and in form, from those that are implied by the standard model of blame-responsibilities. The former put questions of consequences to the fore, the latter questions of character to the fore.

It is the argument of this chapter that these differences matter. In it, I shall be trying to show that my model of task-responsibility makes better sense of the way we apportion differential responsibilities than does the standard, character-based model of blame-responsibility.

I

Consider this distressingly realistic pair of scenarios. In the first, a terrorist gang has exploded a bomb hidden in some rubbish at the back of a busy store, killing several innocent shoppers and injuring several more. How should we apportion responsibilities for this damage? Intuitively, and after considerable reflection as well, we would ordinarily say something like this. The terrorist who masterminded the attack bears greatest responsibility for its consequences. The terrorist who actually detonated the bomb bears a greater responsibility

than the member of the gang who merely planted the bomb in the store and did nothing more. All of those agents bear a greater responsibility than the store detective whose perfunctory search of the premises that morning failed to uncover the bomb, but his responsibility is greater than that of the janitor who failed to clear the rubbish the night before. All of them, however, bear greater responsibility than the politicians and the electorate whose decisions gave the terrorist gang a justifiable grievance.

Or, for a second scenario, imagine a particularly messy divorce that has so alienated a mother from the child left in her custody that she neglects it and it dies in consequence. How should we apportion responsibilities for this death? Intuitively, and after considerable reflection as well, we would ordinarily say something like this. The brunt of responsibility for the child's death must fall upon the neglectful mother. But a substantial, though lesser, measure of responsibility falls upon the ex-husband who, had he exercised court-awarded visiting rights, could have prevented the tragedy. A lesser, though nonetheless substantial, measure of responsibility falls to the Public Health Visitor, who should have been calling at the family's house to check on the health of infants in general. And some modicum of responsibility falls to all those (the child's neighbors, grandparents, etc.) who might reasonably have been expected to observe the child's distress and report it to the proper authorities.

What is revealed by these cases, and hundreds of others like them, is that we ordinarily want to make fairly fine-grained judgments about the relative degrees of responsibility borne by various actors in the sequence of events that conspired to produce any particular state of affairs. But what sense can be made of these judgments? On what basis can we reasonably apportion responsibility?

II

Advocates of the model of blame-responsibility offer a bifurcated response to such questions. First and foremost, anyone possessed of *mens rea* – of evil intent – is criminally guilty, and bears proportionately greater responsibility than anyone in whom this element of intentionality is missing. Second, and very much secondarily, anyone guilty of negligence bears some responsibility; just how much depends upon the magnitude of the harm and the probability that the negligent agent should have foreseen its occurring, given his actions.

At first blush, this account might seem broadly right. Surely it is correct to say that, in the case of the bomb blast, all the terrorists (who harbor evil intent) come in for greater shares of responsibility than do watchmen, janitors or politicians (who are guilty of mere negligence).

And it is probably also right to say that everyone whom we want to hold responsible to any extent at all, in that case or any other, is guilty (at the very least) of negligence of some sort of another.

Upon closer analysis, however, this initial plausibility fades. Questions arise concerning the adequacy of this model on three levels: its account of first-class, intentions-based responsibilities; of second-class, negligence-based responsibilities; and of the relationship between these two classes of responsibilities. On none of these points is the model of character-based blame-responsibility wholly convincing.

A

At the higher level – that of intentions-based responsibilities – the main problem is that this model of blame-responsibility does not allow us to make any very fine distinctions between different people's differential responsibilities. Dealing strictly in terms of intentions, we can make only the grossest distinction between those who were party to the terrorists' conspiracy and those who were not. Suppose the terrorist gang held long and intense discussions about its plan of action and the anticipated consequences. Suppose every member of the gang agreed to the plan, not just in the knowledge that innocent shoppers would die, but actually welcoming that result as part of some larger political program. Then in looking at intentions alone, we would be unable to distinguish between the responsibility borne by the terrorist who actually detonated the bomb and that borne by the one who merely planted it.

Now, perhaps that is the right answer after all. Maybe we were simply wrong in attempting to impose any distinction here in the first place. But there is something undeniably disconcerting about a theory that obliges us to say that there is no moral difference whatsoever between central actors (those who masterminded the plot or pushed the detonator) and bit players (those merely buying explosives or acting as decoys while they were planted), just because they were all coconspirators sharing the same intention. And there is something odder still in saying that the same responsibility – *exactly* the same – is borne by someone who, although participating wholeheartedly (but not figuring especially centrally) in the planning sessions and thus sharing the group's intentions, fell ill and could take no part whatsoever in carrying out the plot. A theory that would hold equally responsible those who played virtually no role in the murders and those who were most instrumental in it, just on account of some shared intentions, is a very queer doctrine indeed.

What has caused all this trouble is something deep within the notion of "intention." Intentionality itself does not admit of degrees. On

the most plausible analysis, some particular outcome is said to be intended by you if and only if that outcome is necessarily part of what it is to accomplish your desired ends.¹ Thus, the death of many innocent civilians must be regarded as an intended (because intrinsic) part of the American plan in dropping the atomic bomb on Hiroshima, rather than as an incidental by-product of American intentions to bring the war to a speedy conclusion. And this can best be seen by noticing that if all noncombatants had been evacuated from Hiroshima before the bombing, then the Americans would have had to choose as their target another city that was still functioning as a "normal" city. (Otherwise, the bombing would amount to the sort of mere "demonstration" that American planners explicitly rejected.) An effect is properly regarded as a foreseen but unintended by-product of our intentional action only if, were the effect in question somehow removed, we would still want to persist in the same plan of action. If that is not the case, then that effect must be regarded as an intrinsic part of our intentions.

Understood that way, an outcome is either intended or not. The question is not *how much* you would have to change your plans if that consequence were to be avoided; the question is merely *whether* your plans would have to be changed. On its face, then, this analysis seems to offer no way to accommodate "degrees" into talk of intentionality.² Yet without that, it seems impossible to accommodate our desire to talk about "degrees of responsibility." We would apparently be unable to "apportion" responsibility in any meaningful sense. We would just have to assign whole masses of people to one category or another. It would be impossible to say that one person is "more responsible," or "bears a greater responsibility," for an outcome than a variety of other people who may share the same intention but who performed a variety of really quite different acts.

Several maneuvers are available to advocates of intentions-based accounts of responsibility in trying to evade these criticisms. One of them is to point to a variety of conditions that diminish one's responsibility for an action by qualifying the claim that the action was indeed

¹ Dworkin (1985, p. 453), borrowing from J. Thomson (1977, chap. 19).

² Similarly, if an agent's "intent" is defined in Davidson's (1980) terms of the "descriptions" under which the agent has chosen the action, then any particular aspect of the action must be either intended (i.e., figure in the agent's own act-description, under which the agent chose the act) or else not intended (i.e., not so figure). Talk of "degrees of intentionality" must, in these terms, mean something like this: "This act-description is relatively nearer to (or further from) the agent's own." But in these terms there is no room for such a metric. Either the act-description is the agent's own (in which case the act was fully intended under that description), or else it was not (in which case it was not at all intended under that description).

performed intentionally. The question is simply whether qualifications such as these can fully account for the systematic variation in the responsibilities of various members of our hypothetical terrorist gang. The answer to that question seems certainly to be "no." While in any particular instance it may be true that the bomb-planter has better excuses than the bomb-detonator, there is no reason to believe that that will always be the case. Indeed, sometimes the opposite is bound to be the case – sometimes it will be the bomb-detonator who is participating under duress, to save his children who are being held hostage by the terrorist gang. Talk of excusing, mitigating or extenuating circumstances seems constitutionally incapable of accounting for the *systematically* greater responsibility of the one party. By its very nature, all such talk must deal in terms of particular features of particular circumstances, and these are bound to vary.

A second evasive maneuver available to advocates of intentions-based accounts of responsibility is to suggest that those with lesser responsibility have in effect "copped a plea," pleading guilty to a lesser charge in exchange for being let off the more serious one. The strategy here is to subdivide the plan and assign responsibility for different parts of its formulation and execution to different agents. Thus, the bomb-planter might be thought guilty merely of "conspiracy to murder" or of "endangering life," whereas the bomb-detonator might be thought guilty of murder in the first degree.³ But if we stick exclusively to the story about intentions, and if intentions are conceived in the terms described here, then you would have to plead *fully* guilty to the lesser charge and be released *entirely* from responsibility on the graver one. That is to say, any gradations of responsibility must be located in gradations of more or less serious offenses, not in differential degrees of responsibility for the same offense. Perhaps it might be logically possible to invent enough distinct offenses to make all the fine-grained distinctions we think we should in our system of responsibility. But it seems empirically unlikely that there are enough already in the statute books or the equivalent moral registers. The decoy or purchaser of explosives bears greater responsibility than the financial backers or passive members of the gang, for example, but it is hard to think how to describe all those differential responsibilities in terms of responsibilities for different moral or legal offenses. Hence it seems that we are holding them differentially responsible for the same thing, rather than holding them responsible for different things.

A third and final evasive maneuver available to advocates of intentions-based models of blame-responsibility is to reformulate the notion of "intentions." On this revised account, what each participant in

³ Feinberg 1970, pp. 244-6.

the terrorist conspiracy intends – but all that any of them intends – is to play one's *own* part in the plan. Crucially, on this account, none of them (with the possible exception of the mastermind himself) would truly intend the outcome of their combined actions. Differential responsibilities would then follow straightforwardly. Those who intend lesser contributions are less responsible for the overall outcome than those who intend greater contributions. This might be a reasonable account of responsibilities for *accidents*, where the intended actions of various agents interact in ways that none of them intends in order to produce the outcome that all regard as unfortunate. It is not, however, a plausible account of conspiratorial intentions. There, the interactions between their intentional actions are deliberate, not accidental or incidental. The intentions of each conspirator "track" the larger goal of the conspiracy as a whole.⁴ If any one of the conspirators alters in any way his intention to participate in the conspiracy, all the other conspirators would (if they know of his altered intentions) consider what alterations are therefore required in their own intentions in order to achieve the aims of the conspiracy. In light of this, no true coconspirator could plausibly claim to be performing intentionally just his own component of the overall plan. Rather, each intends the whole.

B

The second half of the model of blame-responsibility assigns responsibilities of a weaker sort to people on the basis of their negligence. The problem with this half of the model is not that it lacks an account of differential responsibilities, as does the other intentions-based half. Clearly, different people do end up with different degrees of negligence-based responsibility. The question is whether the right people end up being assigned the right amounts.

The extent of negligence-based blame-responsibility varies explicitly with (1) the magnitude of the harm that occurs and (2) the foreseeable probability of its occurring, given the negligent party's actions or omissions. Thus, it is easy to see how one person can be said to be more negligent (and hence more responsible, in that sense) than another. On the analysis here in view, that amounts merely to saying that either the magnitude or the probability or both were greater in his case than in the other's case.

Sometimes that will lead to results that are broadly consistent with our intuitive allocation of responsibilities, but often it will not. In the case of the terrorist attack, we are inclined to hold the security guard

⁴ Nozick 1981, pp. 317–26.

whose search failed to turn up the bomb more responsible for the ensuing mayhem than we are the janitor who merely failed to collect on time the rubbish in which the bomb was hidden, and that may be because security guards hired to search for bombs are more likely to come across them than are janitors.

Further down the ladder of responsibility, however, doubts about the model's adequacy really begin to mushroom. Consider the responsibilities of politicians and electorates who vote for policies that are sure to provoke terrorist bombing campaigns. We are inclined to say, both at first blush and upon reflection as well, that they bear less responsibility for the deaths than the guard or the janitor. But their actions indisputably put more people more at risk of death from terrorist bombs than the actions of either of those store employees. Even though both the magnitude of harm and the probability of harm increase more as a result of their actions, however, we want to hold them less responsible. The negligence model, analyzing blame-responsibility purely as an amalgam of those two factors, is at a loss to explain why.

Similarly in the case of child-neglect, we are inclined to hold the Health Visitor more responsible for the child's death, even though that official seems less "negligent" (on the magnitude/probability calculus of negligence) than the child's neighbors. The neighbors, after all, could observe at close quarters a consistent pattern of neglect over a fairly protracted period. Surely there was a higher probability of a more disastrous outcome resulting from their failure to report the case to the proper authorities than from the Health Visitor's failure to perform the sort of cursory inspection that is all that the law requires. Yet we hold the Health Visitor more responsible.

The upshot seems to be that, while the model of negligence-based blame-responsibility is indeed scalar, it operates on the wrong scale. How negligent one has been, and how much blame-responsibility one will therefore have to bear for the outcome, is analyzed within that model in terms of how "irresponsibly" one has behaved. Irresponsible behavior is analyzed in terms of "thoughtlessness," with how much thought one should have given some action being assessed in terms of the stakes involved, their magnitude and the foreseeable probability that the loss will indeed be incurred. The examples just offered suggest that, even in cases of pure negligence (where there is no hint of evil intent), responsibilities are in fact apportioned in ways other than the probability/magnitude model would suggest. It is not just reflection upon people's characters – not just a disdain for "thoughtlessness" (of that sort, anyway) – that guides our judgments about their responsibilities.

C

Finally, there are problems with the account offered by the blame-responsibility model of the relationship between these two levels of responsibility. Both Kantian philosophy and the legal analogy seem to suggest that intentions-based responsibilities are an order of magnitude more serious than negligence-based responsibilities. The relationship between the two, on the blame-responsibility model, seems to be virtually lexicographical. But anything like a lexicographical ranking badly overstates the extent of the priority which intentions enjoy in our thinking about responsibilities. At the margins, we can easily conceive of cases of really gross negligence that we would consider more blameworthy than intentional minor contributions to a harmful outcome.³

For an example, consider a security guard who did not merely overlook the bomb while doing his ordinary morning rounds, but was instead appallingly negligent. Suppose this was an important anniversary that the terrorists regularly celebrated with bomb attacks; suppose this particular store was a favorite target, having been bombed often before; suppose the previous bombs had all been planted in just the same sort of rubbish piles in just such a corridor. Suppose the security guard in question knew all this. Yet knowing this, and knowing that he will be the only guard to patrol the corridor, he gives the rubbish in which the bomb is hidden only a cursory kick and reports back that there is no bomb there. Finally, suppose that it is a case of pure, unmitigated negligence – the guard is neither a secret sympathizer with the terrorists, nor is he prevented from doing his duty properly by paralyzing fear or threats against the lives of his children or such like. Although the guard is merely negligent in the performance of his duty, he is so grossly negligent that the responsibility he bears for the shoppers' deaths is at least as great as that borne by some other members of the terrorist gang who made only very minor contributions (e.g., acting as decoy while the bomb was planted) to the success of the bomb attack.

One such case is enough to show that intentions-based responsibilities do not enjoy lexicographical priority over negligence-based ones. There may be some scope for argument over whether this reversal of the "natural" order is merely a marginal phenomenon. I am inclined to believe that the margins within which it occurs are really rather wide. But there is no need to enter into that argument here.

³ Indeed, "gross negligence" can sometimes give rise to *criminal* liability (Hart 1968, p. 137).

Advocates of models of blame-responsibility, if they admit that this occurs even at the margins, are then obliged to produce a scalar for intentions-based responsibilities, and to produce some mechanism for mapping this scale onto the scalar for negligence-based responsibilities to tell us how to trade the one off for the other. Only a tight lexicographical ranking of intentions-based over negligence-based responsibilities would have allowed them to dispense with either (much less both) of these elements in their model. But any such tight lexicographical ranking fails to do justice to our ordinary judgments (intuitive and considered ones, alike) about how to apportion responsibilities in some important cases.

III

All those are problems that arise within the model of blame-responsibility. There, assigning responsibilities is seen as being essentially a matter of assigning blame. On the model of task-responsibility, in contrast, assigning responsibility amounts essentially to assigning duties and jobs. Different people have different responsibilities, *ex ante*, because they are allocated different duties and tasks. And people bear differential *ex post* responsibilities for outcomes, on this account, depending on the role that they played or should have played, pursuant to those *ex ante* task-responsibilities, in producing or averting those outcomes.⁶

There are, broadly speaking, two ways in which people might be said to bear different task-responsibilities. In the first, they will have been assigned responsibility for different tasks. In the second, they will have been assigned different responsibilities *vis-à-vis* the same task. An important example of the latter is the way in which we assign some people (e.g., godparents) secondary, backup responsibilities for stepping in and doing certain things (e.g., raising their godchildren) when those with primary responsibility (e.g., their natural parents) fail to discharge it.

I shall say very little about the bases upon which these task-

⁶ Failure to discharge a task-responsibility might consist in either an act or an omission, depending upon what the task-responsibility required of the agent. Even those who attach great moral importance to the distinction between acts and omissions concede there is not necessarily any moral difference between them where the agent is under some *special* responsibility, though. A sleeping sentry is as blameworthy for any harm that results from his negatively omitting to discharge his responsibility as a drunk driver for any harm that results from his actively violating his. See Goodin 1985c, chap. 2, and references therein.

responsibilities get assigned to the particular people they do.⁷ No doubt part of the story – no doubt a large part of it – has to do with people's differential capacities for performing the tasks and duties at issue. Perhaps some people are peculiarly well situated to act, or to know if and when action might be necessary. Or perhaps they are particularly blessed with the peculiar sorts of resources (be they material or affective) which the task requires. Or perhaps, in cases where it is crucial to the success of the enterprise that some people act and that others stay well out of their way, they possess some peculiar attribute that makes it "obvious" to themselves and to others that they are the ones who should act. All of these are essentially consequentialistic considerations, assigning task-responsibilities to people on the grounds that they are particularly able to discharge them. But there may also be other nonconsequentialistic considerations at work in the allocation of task-responsibilities. Whether or not there are is a question that, for present purposes, can safely be left open.

My main concern here is instead to show that the model of task-responsibility can offer a plausible account of the way in which we do, in practice, apportion responsibilities among different people. Specifically, the model of task-responsibility can successfully avoid the pitfalls found, in Section II, with the model of blame-responsibility.

The first problem, discussed in Section IIA, is that the blame-responsibility model, because of its narrow focus upon people's intentions in allocating top-level responsibilities, cannot account for the differential responsibilities borne by people sharing identical intentions but taking very different actions. On the model of task-responsibility, there is no such difficulty in accounting for the differential responsibilities of the various members of the terrorist gang (the mastermind, the detonator, the decoy, the passive member, etc.). The task-responsibility analysis would start from the proposition that each of these people had, *ex ante*, a duty (understood here as a negative task responsibility) to make sure that they did not cause harm to others. From there, it would go on to observe that all these various agents have failed, more or less badly, to discharge that task. How "badly" they have failed depends, in turn, on just how large a causal contribution each has made to the proscribed outcome. The relative strength of each party's task-responsibility for the terrorist bombing thus depends upon the relative importance of the role each played in the overall drama. Lesser degrees of responsibility fall to those (such as passive gang members and decoys) who played what were only very minor roles, in the sense that the basic plot – the

⁷ My own broadly utilitarian views on these issues are elaborated in Goodin 1985c, chap. 5.

overall outcome – would have been little changed if they had been omitted altogether. Greater degrees of responsibility fall to those (such as the mastermind and bomb-detonator) who played major, crucial roles in the drama.^a

Second, the model of blame-responsibility runs into difficulties in analyzing lower-level responsibilities purely in terms of negligence, understood as probable harm. To revert to the example of Section IIB, we are often inclined to hold Health Visitors more responsible for failing to report cases of child neglect that we are the child's neighbors, even if in any particular instance the neighbor's inaction is more probable to cause more harm. Why the Health Visitor should be held more responsible in this way is a mystery on the model of blame-responsibility's account of negligence. But it makes perfectly good sense on the model of task-responsibility. On that account, the essence of negligence is not probable harm, but rather the shirking of one's duties. It is the job of the Health Visitor, in a way it is not the job of the neighbors, to detect and report cases of child neglect. If the Health Visitor fails in that duty, the Health Visitor therefore bears more responsibility for the subsequent death.

Third, the model of blame-responsibility suffers from problems, discussed in Section IIC, arising from the sharp disjunction in its treatment of higher (intentions-based) and lower (negligence-based) levels of responsibility. In the model of task-responsibility, everything depends upon how important a task you have been assigned and upon how well or badly you perform it. Thus, it is perfectly capable of saying – in a way that the model of blame-responsibility cannot, or can only after considerable twisting and turning – that the grossly negligent guard is more responsible for the innocent shoppers' deaths than is the financier or the perfectly passive member of the terrorist gang. There is no guarantee that there will always be a perfectly continuous measure of role-importance in every instance, of course. But at least there is no built-in guarantee of the sharp and systematic discontinuity that characterizes the model of blame-responsibility.

Different people have different responsibilities, on the model of task-responsibility, most fundamentally because different people have been assigned different tasks. Blame for failing to discharge those task-responsibilities may be greater or lesser, depending upon how large a role they require one to play in producing the outcome in question. Those with minor roles bear less responsibility than those with major roles.

^a If the success of bombing attacks usually depends mostly on successful coordination of activities of several coconspirators, the mastermind who performs that coordination role bears the greatest share of the responsibility.

Blame for failing to discharge one's task-responsibilities may also be greater or lesser depending upon how clear one's task-responsibilities were at the time one was expected to discharge them. Sometimes one is justifiably uncertain whether or not something is one's own responsibility – one's task, duty or job – at all. Other times one is justifiably uncertain as to what particular actions are required in order to produce some particular outcome that one is responsible for making sure is produced. If the *ex ante* perspective of the task-responsibility model is the ruling one, then one's responsibility must be less the less clear it was at the time what one's task-responsibilities were or what they required one to do. And surely this is the right way to think about responsibility. A security guard who failed to search the corridor in which the terrorist bomb was hidden would surely have more success in defending his negligence by saying that he justifiably thought that someone else was supposed to be searching that corridor (the classic excuse for negligence on the task-responsibility model) than he would by saying that he justifiably thought it highly improbable that any very large bomb could have been hidden there (the classic excuse for negligence on the blame-responsibility model).

IV

There is one final piece of evidence in favor of the task versus blame model of responsibility. It is just this. We can say quite confidently that the terrorist setting off the bomb is more responsible than the terrorist planting it for the deaths of the innocent shoppers, and that both are more responsible than the security guard or the janitor. By the same token, we can say quite confidently that the neglectful mother is more responsible than the slack Health Visitor for the starvation of the infant. But we are hard-pressed to say, in any very precise way at all, exactly *how much* more responsible any of these agents is than the others.

Now, there is no reason on either model to suppose that these ratios should be particularly easy to calculate. So the fact that we find the calculations difficult is not necessarily evidence, in and of itself, in support of either model.

What does count as decisive evidence in favor of the task-responsibility model is not that we find precise calculations of relative responsibilities difficult, but rather that we find them *dispensable* for such a broad range of practical affairs. For the *ex ante* purposes guiding my choice of actions, I merely need to know whether it is more my responsibility than yours to see to it that some particular task gets done. How much more my responsibility than yours it is, from this *ex ante*,

Apportioning responsibilities

action-guiding perspective, does not matter. Just so long as it is clearly more my task than yours, the allocation of task-responsibilities will be clear enough. The fact that we ordinarily seem so satisfied with ordinal rankings of responsibilities, and so indifferent to the absence of interval-level measures of their relative strengths, thus powerfully suggests that ours is first and foremost a concern with that kind of responsibility that assigns tasks rather than fixes blame.⁹

* I am grateful for comments on this chapter from Diane Gibson, Martin Hollis, Philip Pettit and Jeremy Waldron.

PART IV

Shaping public policies

Respecting and overriding preferences

Liberalism and the best-judge principle

Utilitarianism, in its various forms, is an essentially preference-respecting, want-regarding morality.¹ In its insistent refusal to brook any perfectionist notion of The Good which is good independently of anyone ever thinking it to be good, utilitarianism just echoes the more general Enlightenment liberalism from which it sprang. Arguably the defining feature of modern liberalism is its value-neutrality, its agnosticism as regards matters of ultimate personal value.² In liberalism more generally, as in utilitarianism more narrowly, each individual is deemed to be the "best" (indeed, ultimately the "only true") judge of his own interests.

That claim can be construed in various different ways, however, depending on whether we talk about what people actually do or sometime will or counterfactually would want. Therein lies the core of the controversy between liberal utilitarianism and its liberal-libertarian cousins. Libertarianism is a radically antipaternalistic doctrine, taking whatever preferences people presently express as absolutely veridical. At root, utilitarianism is equally respectful of people's preferences. Utilitarians, however, are more sensitive to nuances within people's preferences and are more sensitive, in consequence, to possibilities for finding some warrant in people's own preferences for overriding some of the preferences they presently express.

What is at dispute between interventionist utilitarians and laissez-faire libertarians is, thus, merely the sense in which people ought to be said to be the best judges of their own interests. In this chapter, I identify four versions of the best-judge argument. Libertarians opt for strong versions of the claim. They urge that principle as an axiom of liberal value systems, as the unexamined and unexaminable premise

¹ Barry 1965.

² R. Dworkin 1978. Raz 1986. Goodin and Reeve 1989.

on which all the rest must be built. But even axioms have to be coherent. Their injunctions have to be consistent and comprehensible if they are to serve to guide action at all. Furthermore, although axioms cannot strictly be "derived" from anything else in the system of values, they nonetheless have to be "motivated." There has to be some reason for adopting one set of axioms rather than some other, and the reasons for adopting that set have to be stronger on balance than the countervailing reasons for not adopting it. Libertarians' stronger non-empirical claims on behalf of the best-judge principle fail both these tests.

What is left is the utilitarians' weaker, empirical version of the best-judge claim. What is only contingently true is sometimes contingently untrue; and as I shall show in Section IV, there are three broad classes of cases in which familiar psychological dynamics work to ensure that people are not particularly good judges of their own interests and in which broadly utilitarian public officials might be expected to be better judges. The systematic failure of the best-judge principle in such cases undermines laissez-faire policy prescriptions of a liberal-libertarian sort and paves the way for interventionist policy of a liberal-utilitarian sort.

I. THE ONLY-JUDGE PRINCIPLE

The boldest version of the best-judge principle holds that an individual is not only the best judge of his own interests; he is the *only* judge. He is the sole source of authority on such matters. His interests are just what he says they are, no more and no less.³

What others are doing when trying to judge his interests is, on this account, merely trying to predict his own judgments of them. Just as with court-watchers trying to predict a judge's verdict, such predictions may be more-or-less accurate, more-or-less well founded, and so on. But predicting someone's judgment of his own interests is, on the "only-judge principle," rather like a legal realist's prediction of the court's decision.⁴ In neither case can your prediction be based on any perception of some independently right answer, which you discover and then predict that the agent himself will also perceive and act upon. On such models as these, there is no right answer, apart from the answer that the agent himself ultimately delivers. His judgment, in both cases, is strictly veridical, in and of itself.

The "only-judge principle" draws such support as it enjoys from a

³ "No one but the person himself can judge" (J. S. Mill [1859] 1975b, chap. 5, p. 118; cf. von Mises 1949, p. 19).

⁴ Llewellyn 1960. Holmes 1897.

certain perversion of the "privileged-access" argument. Each of us, it is often said, has privileged access to his own mental states, conspicuously among them thoughts, feelings and emotions. Much though others may empathize with me – imagining themselves in my place, feeling pain in my pain and taking pleasure in my pleasure – they are in the final analysis only feeling feelings *like* mine. They cannot experience my feelings without putting themselves in my place quite literally, occupying my body and my mind as well.⁵ So, too, even with mere thoughts. You may share my thoughts, in the sense of thinking thoughts *like* mine, that is, with the same propositional content. But you cannot think my thoughts, literally, without occupying my brain (and, if it is a separate entity, my mind).

All that once seemed reasonably noncontentious. It no longer does. It is now clear that privileged-access arguments will need considerable recasting, if they are to be salvaged at all.⁶ For present purposes however, I propose to leave all that to one side. My objection here is not going to be to privileged-access arguments in general, but merely to the way they are used in support of the only-judge principle.

My objection, in essence, is that the privileged-access argument and the only-judge principle refer to fundamentally different things. The former refers to the inaccessibility of other people's *minds*; the latter refers to the impossibility of our judging their *interests*. However intimate the connection between people's interests and their mental states, it is clear that they are not one and the same thing. People can have privileged access to their mental states without being the sole judges of interests.

That simply must be so. Otherwise it would be impossible for statements about people's interests to serve the social function that such statements are meant, by those people themselves, to serve – viz., to guide social action. In order to serve as social guides, statements about interests must be interpersonally intelligible. (We cannot be guided by what we cannot comprehend.) Yet if the strong form of the privileged-access argument embodied in the only-judge principle is true, we have no way of rendering intelligible others' statements about their own interests.

What makes people the "only judge," after all, is that no one else can get inside their heads and no one can know their interests without so doing. Yet if we cannot know their interests without getting inside their heads, by the same token we cannot know their interests on the

⁵ "Our senses will never inform us of what [another] suffers. They never . . . can carry us beyond our own person" (Smith 1790, pt. 1, sec. 1, chap. 1; cf. B. Williams 1973b, chap. 1). New welfare economists object to interpersonal utility comparisons on similar grounds (Jevons 1911; Robbins 1932, pp. 122–5; 1938).

⁶ Davidson 1987; Heil 1988.

basis of secondhand reports of what is going on inside their heads. (After all, those secondhand reports do not themselves literally take us inside their heads; and on the privileged-access interpretation of the only-judge principle, that is what is crucial for knowing a person's interests.) Hence, the selfsame thing that on the privileged-access/only-judge principle makes firsthand judgments of others' interests impossible renders those others' own secondhand reports of them equally unintelligible.⁷

Now that is nonsense. Others' statements about their interests (thoughts, feelings, emotions and so on) *are* perfectly intelligible to us. But insofar as they are, then whatever way we have of rendering others' interest-statements intelligible – and hence action-guiding – will also allow us to “judge” their interests, in some strong sense, for ourselves. That is to say, it does not merely allow us to make more accurate predictions of what they will say their interests are; it also provides some leverage for allowing us to say what they *should* say they are. What is at issue here is not whether we are right more often in our judgments than the people themselves. It may well be that we are not. The person concerned is still the *best* judge of his own interests. My only point here is that he is not the *only* judge. We have some capacity for judging his interests ourselves and this is the same capacity as is required to make intelligible his own interest-statements.

Statements of interests entail *claims*, not just demands. To say “I want *x*” is a demand. We do not need to know why, only that you want *x*, to comprehend fully the nature of the proposition and (if we are in the business of want-satisfaction) to act upon it. To say “It is in my interests that I should have *x*” is to say something about the *reasonableness* of a want. It explains, rationalizes, justifies a want, in the way that a statement of brute longing does not. It points to some further fact that renders the want comprehensible, understandable, reasonable to others. It is this reasonableness to which we – and especially we liberals – are responding when striving to honor people's interest-claims. But if the function of interest-claims is to *justify* our demands to others – to provide some interpersonally intelligible reason for our having those demands met – then others must have some mechanisms whereby they can, in effect, judge your own interests. The same thing that makes it possible for others to understand your interest-statements makes it possible for them to judge your interests, well or badly, for themselves.

That is partly because of the need to *validate* interest-claims; that is, to make sure that they are not false, trumped-up claims, before honoring them. But that is only a small part of the story. Mostly, inter-

⁷ Davidson 1986.

personal intelligibility of interest-statements is necessary because to understand the claims at all I need to understand more than *just* that we both have wants (or needs or longings).⁸ To understand how that gives rise to a claim to *x* – a claim that is rationally compelling for me to honor – I must understand that *x* (e.g., food) matters in the same way to him as to me, or that *x* means to him what *y* means to me.⁹ In short, I need to know not only *that* he wants *x* but also *why* he wants *x*. In coming fully to comprehend why, I must end up mirroring his mental processes, judging his interest in *x* as he judges it himself. All of which means, of course, that I end up judging his interests along with him, if I am to act on his interest-statement (rather than just his demand-statement) at all. But if I judge his interests along with him, then he is not the *sole* judge at all.

Indeed, the processes I go through in judging his interests along with him, in making sense of his interest-claims in that way, I could equally well have gone through in judging his interests without him. What I do to comprehend his claim after he has made it could equally well have been done before he has pronounced on the subject of his interests at all. Not only is the individual himself not the sole judge, in the sense that for interest-statements to serve their social function someone else judges them alongside him; he may not even be an indispensable party to the judging after all. Others can gather the information they need about the role of hiking in Jack Smart's life to conclude that it is in his interests to have easy access to open country, without having had him first put the pieces together for himself and so pronounce it in his interests.

II. THE SOLE-LIABILITY PRINCIPLE

There is a pair of variations on the "only-judge" theme that should be considered next, albeit briefly. Both concede that the only-judge claim is itself too bold. Both allow that a person is not the sole judge, nor perhaps even the best judge, of his own interests. But both go on to say that, while it is perfectly possible for others to judge his interests, and perhaps better than he can himself, his own determination of those interests ought nevertheless to be the only one that counts.

The first such argument says the reason is this. We ought to respect a person's own assessment of his interests, not because he is the sole judge of those interests, but because he alone will suffer the consequences of judging them well or badly.¹⁰ While renouncing the bolder

⁸ Cf. C. C. W. Taylor 1969; Norman 1971, chap. 3; Watt 1972.

⁹ Scanlon 1975.

¹⁰ As J. S. Mill ([1859] 1975b, p. 102) says in only a slightly different context, "the decision ought to rest with those who are to abide the consequences."

"privileged-access" claims of the only-judge principle, this "sole-liability principle" crucially retains the essential liberal policy prescription. Others may judge your interests, perhaps from time to time better than you can yourself. But given that you alone will suffer, it is none of anyone else's business to interfere, imposing their superior vision of your interests in place of your own.

There are some circumstances in which this – and more – is true. Consider in this connection Locke's argument in his *Letter Concerning Toleration* against compelling people to worship the One True God. Such worship would secure their salvation but only if based on sincere belief on the part of the worshipper. Forcing people to go through the motions would simply not secure the end in view.¹¹ Put in terms of interests, the point is this. There are some things that others can see would be in our interests to pursue, but from which benefits would follow only if we ourselves saw that they are in our interests to pursue. Hence, acting upon (still more, being forced to act upon) others' perceptions of our interests would fail to secure such goods at all. Roughly speaking, any case where the actor's intentions matter would be susceptible to this sort of argument. "My mother/lawyer/confessor told me it would be best if we made up and became friends again" is not a very convincing basis for a friendship. Similarly with all manner of personal relationships, an agent's interests in them can be pursued only if he himself takes (or, minimally, is thought to take) an interest in them.

Such interests as this – where the agent can serve his interest in x only if he sees himself as having an interest in x – are only a small subset of all an agent's interests, however. For that subset, the sole-liability principle and the policy conclusions that follow from it are powerful indeed. But the more typical case is one in which the agent's own perceptions are in no way constitutive of his interest-satisfaction. In those cases, the sole-liability principle would seem to amount to little more than a shaky empirical proposition. Misperceiving your interests *can* sometimes – perhaps even standardly – have important consequences for others. Suppose the Soviets had implausibly believed a few years ago that the United States would suffer a "window of vulnerability" during which the USSR could launch a broadly successful preemptive strike on the United States. Suppose that, based on this misperception of their interests, they launched an all-out nuclear attack on the United States, which was not successful in preempting completely an American response and they suffered substantial losses as well. Now, certainly the Soviets would have suffered from their misperception but they would hardly have been the

¹¹ Locke [1689] 1946.

only ones to have suffered. Or suppose that you pull out to pass a slow-moving tractor misperceiving the distance to the oncoming car. You suffer as a result of your misperception but the occupant of the car you crash into also suffers. It is simply not the case that you are always the only one to suffer as a result of your misperceptions of your interests. Where the sole-liability principle is merely pointing to an empirical claim rather than an analytic one, it is empirically false as often as not.

The way to save the claim is, presumably, to rephrase it. A revised version might hold that "people ought to be taken to be the only judge of their own interests *only if* they alone will suffer the consequences." The "only if" in this formulation replaces "because" in the bolder version. This, however, amounts to saving the principle by denying its application to very many cases. First we have to establish what everyone's interests objectively are in the various policy choices before everyone. Then we have to establish who suffers as a result of each of the choices that everyone is inclined to make. Only then can we say if the sole-liability principle should be applied. If no one except the agent himself will suffer (directly or indirectly, in the short or long term) from what the agent proposes to do, we should let him proceed. It is an empirical question how often that principle happens to be applicable. It is an article of faith in the worst sense – that is, scarcely credible – that it is true often enough to form a very important part of our practical instructions to social policy-makers.

III. THE SOLE-RESPONSIBILITY PRINCIPLE

A second variant of the only-judge principle is one that concedes that others can judge your interests – perhaps, indeed, better than you yourself can – but again concludes that you ought to be left to judge your interests for yourself anyway. With the "sole-liability principle," the rationale for doing that was that you alone would suffer if you got them wrong. Here, the rationale is that you will have only yourself to *blame* if you do so. Judging your interests is your responsibility and yours alone. Making your own choices and living with the consequences is – on at least one account – what autonomy and self-government are all about. Let us call this the "sole-responsibility principle."¹⁴ This runs into none of the problems of an empirical sort that plagued the sole-liability principle. There is no claim that the

¹⁴ Maybe that is what J. S. Mill ([1859] 1975b, p. 115, emphasis added) meant when saying that the maxim running throughout *On Liberty* is that "the individual is *not accountable to society* for his actions, in so far as they concern the interests of no one but himself" (see similarly Arneson 1980; Feinberg 1986, chap. 18; Benn 1988, pp. 12–13).

individual is the sole judge because (or where) he alone will suffer if his interests are misjudged. The claim is instead that each individual should "carry the can" for his own errors. If his misperceptions cause harm, to himself or to others, then the responsibility is his.

The principle is susceptible to empirical objections of another kind, however. There is some considerable psychometric evidence that many – perhaps most – individuals have a fairly strong preference for responsibility-avoidance. If offered a choice between a bad outcome for which they might be regarded as being responsible and an even worse outcome for which they might plausibly disclaim responsibility, they will opt for the worse outcome almost every time.¹¹

What emerges from those studies is that ascribing responsibility skews choice. People's judgments vary, depending on whether or not they will be held responsible for the outcomes of their choices. So making people responsible for choosing their interests will, perhaps, alter their choice of interests. If so, they will not be saying what is *really* in their interests.

From this empirical objection arises another, rather more principled one. Notice that, whereas other arguments aim to justify *letting* each individual judge his own interests, this version seems to argue for *making* him do so. It is his responsibility to judge his interests and his alone; any consequences that arise from his misjudgments are, therefore, also his responsibility, and his alone.

Various arguments – couched in terms of autonomy, dignity or self-respect – might justify letting people assume responsibility for their own lives and with it responsibility for judging their own interests for themselves. But none of those principles can easily or obviously be turned to justify making people do so if they prefer not to do so. The basic point here is just this. There are a great many ways of manifesting autonomy, dignity or respect for oneself. Judging one's own interests for oneself might be one way but it is not the only one. If someone wants to pursue another path, there is nothing in those larger principles to justify us in preventing him from doing so.

Similarly, accepting responsibility might plausibly be construed as lying at the core of moral agency but there are lots of things one might accept responsibility for. Judging one's own interests for oneself might be one, but only one. A person can accept responsibility for his actions without accepting responsibility for (much less personally undertaking) every calculation upon which they were based. British cabinet ministers do it all the time. Or at least they used to do so.

¹¹ For psychometric evidence, see Tversky and Kahneman 1981 and Tversky 1981. Similar propositions recur in philosophy (G. Dworkin 1988; Reeve 1990) and policy-making (Weaver 1986; McGraw 1990).

IV. THE BEST-JUDGE PRINCIPLE

The upshot of all this is that the most that can plausibly be made out along these lines is a baldly empirical claim. Usually (not always) each individual is probably (not necessarily) the best judge of his own interests. It is not that others *cannot* judge his interests for reasons to do with his privileged access to his own mental states. It is not that others *should not* presume to judge his interests, because (or where) either liability or responsibility will be his alone. Rather, the rationale for letting each person judge his own interests is that, on balance, we suppose that he is more likely to judge them more nearly correctly than anyone else.¹⁴

The truth of that empirical claim is, however, tightly circumscribed. There are many cases – some familiar, some not so familiar – where that is simply not true. If the empirical claim is all that there is to ground liberal and neoliberal policy prescriptions, then in those cases where the empirical claim proves false we may (and ought to) substitute others' judgment for his own.

One such class of cases arises from problems akin to "weakness of will." In the classic case, that refers to someone failing to hold to his resolution to pursue certain goals he has set for himself when it comes to the time to act. Similarly, someone might prove unable to bring himself to do what is necessary to pursue his self-perceived interests when it becomes time to act. Allied with that failure – whether as cause or consequence is not important here – comes a reconceptualization of his interests. The weak-willed person comes to revise his judgments in line with his actions, regarding interests he has not the heart to act upon as not being his interests at all. Yet that is quite false.¹⁵ The strain to reduce cognitive dissonance has simply led him to misstate, in perfectly predictable ways, his own interests.

Similarly, people are notoriously bad at judging interests that hang in any important way upon probability calculations. They are liable to all sorts of biases, in assessing such probabilities, growing out of judgmental shortcuts that are by now well mapped by psychologists and decision theorists.¹⁶ Perhaps most worrisome is the tendency to

¹⁴ "Speaking generally, there is no one so fit to conduct any business, or to determine how or by whom it shall be conducted, as those who are personally interested in it" (J. S. Mill [1859] 1975b, chap. 5, p. 133; see similarly Smith 1790, pt. 2, sec. 2, chap. 2; cf. Feinberg 1986, pp. 57–62).

¹⁵ *Ex hypothesi*, the reason he cannot bring himself to pursue his interests is not just that the costs of doing so are such that, net of costs, the action is not in his interests after all. The weak-willed, like the hot-headed, "act knowingly against their own interest" (Hume 1739, bk. 2, pt. 3, sec. 3).

¹⁶ Kahneman, Slovic, and Tversky, 1982. Goodin 1982b, chap. 8.

assume "it will never happen to me," even when the odds are decisively against that proposition.

Finally, people are not particularly good at judging their long-term future interests. In part, this problem arises from the same source as does the last one: in conjuring up a full and vivid image of what it would be like. Psychologically, it proves enormously difficult to identify either with an accident victim by the roadside or with the doddering old-age pensioner that (with luck) you will become.¹⁷ Insofar as people manifest a wholly irrational bias in favor of present over future interests, intervention can again be justified.¹⁸ They are, in that case, ill suited to judge their further-future interests.

All that leads to the conclusion that there are certain well-defined sets of cases in which people are *not* particularly good judges of their own interests. Generically, these can be characterized as cases involving "adaptive preferences," probabilities or the further future.¹⁹ These are far from uncommon elements. Whenever they are present, there is a case to be made for discounting a person's own judgment of his interests.

Whether anyone else is a better judge of his interests is, of course, a separate issue. However poor the individual himself might be at judging his own interests in such cases, we might nonetheless decide to leave the matter to his own discretion if no one else can or will do a better job of it. Classical liberals have always made much of that point.²⁰ Contemporary libertarians often seem to rely on it almost exclusively.²¹ Notice, however, that in the circumstances here in view there is indeed reason to suppose that public officials can judge individuals' interests better than can those individuals themselves. The general psychological dynamics systematically forcing individuals to misperceive their own interests in such situations do not apply to public officials, who have no personal stake in the matter.²²

¹⁷ Parfit (1984, pt. 3; cf. Nozick 1981, chap. 1) might say that, insofar as you fail to identify with the old-age pensioner, to that extent he is not really "you" and his interests not really "yours," after all. But if the old person is not you, then neither is there any reason to let you judge his interests for him. The less of our lives we can claim as our own, the less scope there will be for applying the liberal-libertarian principle of "letting each judge his own interests," and the more scope there will be for liberal-utilitarian policy-makers promoting policies that are in people's interests whether they (presently) acknowledge it or not.

¹⁸ Goodin 1982a. Parfit 1984, appendix F. People are particularly bad at anticipating future changes in their own preference orderings (Goodin 1982b, chap. 3).

¹⁹ Elster 1983.

²⁰ J. S. Mill 1848, bk. 5, chap. 11, sec. 7; [1859] 1975b, chap. 5, p. 133; [1861] 1975a, chap. 3. See similarly J. Mill [1823] 1992, sec. x (ii).

²¹ Brittan 1988, p. 42.

²² Classical liberals count heavily on people learning from their mistakes (J. Mill [1823] 1992, sec. x [ii]; J. S. Mill [1861] 1975a, chap. 3), but there is no reason to

Just look down the list of circumstances that make individuals misperceive their interests and notice how largely immune from those forces would be public officials judging their interests on their behalf. It is not the public official's face that needs to be saved, through adaptive preference-formation, when weak-willed individuals fail in their resolve. Or, again, public officials, by the nature of their jobs, *must* look to statistical aggregates and plan on the basis of probabilistically expected values. Individuals may tell themselves, one by one, that "it will never happen to me" but public officials know that it will happen to some of them. Or, yet again, individuals may discount future rewards on the grounds that they might not be alive to enjoy them. Whereas individuals are mortal, however, public institutions and the general social interests associated with them are not. That makes public officials take a longer time horizon than do individuals planning their own private lives.³¹

Clearly then, public officials with no personal stake in the matter *can* judge an individual's interests in these sorts of situations better than can the individual himself. The reason is simply that they are exempt from the psychological forces that cause him to misperceive his own interests. Whether public officials *will* judge better than the individual himself is, perhaps, still another issue. "Each man is the only safe guardian of his own rights and interests," John Stuart Mill writes, echoing his father before him.³²

But as his father before him had discovered, requiring democratic accountability of public officials will achieve the best of both worlds. Being judged by voters in the longer term and in a more general way for their superintendence of the citizenry's interests, representatives can do for the people what is truly in their interests but which they would find it psychologically difficult to do for themselves. Being ultimately democratically accountable, they would be unable to do too much that was not in the citizenry's longer-term and more general interests, otherwise they would not win reelection.

No one says these mechanisms of democratic accountability work perfectly and there will always be hard cases at the margins. In any given occasion or on any given set of issues, the peculiar capacities of the people themselves (or the peculiar incapacities of public officials) may be such as to justify leaving the matter with the people themselves to judge. However, the general psychological dynamic in the

suppose that they will and every reason to suppose they will not in the peculiar circumstances here in view.

³¹ Even in their capacity as voters, people take a longer-term view of matters, voting for public works and capital investments that will not pay off until they are long dead (Marglin 1963, p. 98).

³² J. S. Mill [1861] 1975a, chap. 3, p. 187. J. Mill [1823] 1992, sec. x (ii).

sorts of cases I have been discussing is sufficiently well established to lead us to conclude that it ought not be left to them, as a matter of course, in such cases. There, the "best judge" of a person's interests is probably not the person himself but disinterested (or less directly interested) public officials.

V. POLICY IMPLICATIONS

Libertarianism is, thus, a mistake that arises from conflating the two really very different roles that the best-judge principle might play within liberal theory. One is epistemological. How do we know what a person's interests are? Just ask him. The other is political. How can we best further people's interests? Just let them do what they want, so long as they do not interfere with anyone else. The one conclusion follows only very imperfectly from the other, however. What people would do, if left to their own devices, is often at considerable variance with their own accounts of what they want. Thus, *laissez-faire* libertarian conclusions in the sociopolitical realm do not necessarily follow, even if we were to grant the validity of the best-judge principle in the epistemological realm.²⁵

Perhaps the most dramatic case concerns choices for addictive substances – tobacco, for example. The World Health Organization defines addiction in terms of trying to stop consuming the substance and being unable to do so. The fact that people continue consuming the substance should not, in such circumstances, be taken as conclusive evidence that they truly prefer smoking to nonsmoking. On the contrary, their failed attempts to stop should be taken as more conclusive evidence of their deeper preferences in the matter. They should, not incidentally, provide the stronger guidance to public policy-makers contemplating whether or not to pursue policies discouraging future generations of smokers from taking up the habit.²⁶ By the same token, the fact that everyone tries to evade the taxes required to pay for public works projects should not be taken to imply that those roads, dams and sewers ought not to be built. People would truly prefer that all – themselves included – be taxed and such services provided, rather than none being taxed with the consequence that no services are provided. Just watching all of them try to free-ride on the contributions of others, we might never realize that fact. It is undeniably true, nonetheless.

Choices are simply misleading indicators of preference. We must

²⁵ That is just to say that in extracting information from people about their perceptions of their own interests, we should be sensitive to their arguments as well as their conclusions. This theme is elaborated in Chapter 9.

²⁶ Goodin 1989a, b; 1991b.

always look to the reasons underlying choices, rather than to the brute facts of the choices themselves, in surmising real preferences. Doing so, we might find good grounds for concluding that a person's choice of action is wrong, even when judging in terms of that person's own deeper or higher-order preferences.

VI. RECONSTRUCTING LIBERALISM

What I have been arguing for in this chapter amounts to little more than utilitarian-style "interventionist liberalism" of the sort that has been familiar from the days of Joseph Chamberlin and Franklin Roosevelt. Where it is empirically true that people are likely to be the best judges of their own interests, or where for some analytic reason their interests can be attained only by their pursuing them themselves, or where for some contingent reason no one else is likely to pursue them any more effectively, then policies and institutions should allow people to frame and pursue their own interests as they will. But that is not always the case. In those well-defined circumstances in which it is likely to be untrue, public interventions of a broadly utilitarian sort are justifiable on the selfsame want-regarding, preference-respecting grounds as, in other circumstances, underwrite liberals' *laissez-faire* libertarian policy prescriptions.

Precisely what form those interventions should take will naturally vary. Sometimes it is a matter of legal regulation of conduct, requiring people to perform certain actions and to forbear from others, "in their own best interests." Other times, it is a matter of public provision of certain sorts of goods and services, substituting for market provision, on the grounds that people cannot or will not best further their own interests through private choices in free markets. The larger general point is that in such arguably common cases liberal-libertarian *laissez-faire* policy prescriptions ought to be repudiated in favor of interventionist liberalism of a broadly utilitarian, though still ultimately preference-respecting, sort.¹⁷

¹⁷ I am grateful for comments on this chapter from colleagues at the University of York and the Leyden Institute for Law and Public Policy, especially Andreas Kinning, Richard Lindley, Sue Mendus, David Miller and Andy Reeve.

Laundering preferences

Want-regarding moralities like utilitarianism are continually embarrassed by the fact that some preferences are so awfully perverse as to forfeit any right to our respect. Judging states of affairs according to the utilities that they contain alone, we would be unable to distinguish the utility flow that comes from a starving person's being better fed from one that comes from his sadism being indulged.¹ Where such perverse preferences are involved, we are intuitively opposed to ranking social states on the basis of "utility information" alone. Instead, we intuitively suppose we should try to bring "nonutility information" to bear, typically in the form of vested rights guarantees protecting people from the meddlesome (or indeed sadistic) preferences of others.²

The theme of this chapter is that this retreat from utility to rights is premature. If the problem is that preferences sometimes seem "dirty," then surely we should see whether they can somehow be "laundered" before discarding them altogether. The argument of this chapter is that we hesitate to launder preferences only because we are unsure of their fabric.

Recourse to nonutility information seems necessary merely because we work with such an impoverished conception of individual preferences in the first place. For the most part, they are just taken to be an individual's ranking of various social states. Whatever underlies this ordering ordinarily goes undiscussed. But, in truth, there is much more to individual utilities than is captured by simple numbers and rank-orderings.³ "Utility information" can and should be seen to include information about *why* individuals want what they want, about

¹ Sen 1979a, pp. 547–98. See similarly R. Dworkin 1977, pp. 234–8.

² Sen 1970, chaps. 6 and 6*; 1976. Hart [1979] 1983.

³ Lancaster 1966. This is the core of Sen's (1977a, pp. 335–6) own notion of "rational fools."

the *other things* they also want, about the *interconnections between* and *implications of* their various desires, and so on. Obviously, this goes well beyond the sort of information social choice theorists ordinarily ask us to collect – or their models are capable of processing.⁴ But that does not make it nonutility information. The information in question is still very much information about individual utilities.⁵

The ultimate goal of enriching our utility information in this way is to use it to launder people's preferences. The thin theory of individual preferences leaves opponents of want-regarding moralities too easy a task. They need only nod in the direction of some incredibly nasty preferences and say, "There must be something terribly wrong with any principle that requires us to respect *those* preferences." Want-regarding moralities can be spared that sort of peremptory dismissal by showing that there are reasons, internal to preferences themselves, for disregarding some sorts of preferences.

The plan of attack is as follows. Section I argues that "censoring" utility functions is a more adequate response to the concerns that drive us to such nonutility recourses as the ascription of rights. Section II reveals various ways in which laundering preferences could, in principle, be perfectly consistent with respecting preferences construed more broadly. Section III goes on to argue that, in the context of collective decision-making, people are forced to undertake a limited laundering of their own preferences. And Section IV discusses practical devices for such further laundering of people's preferences as Section II might warrant.

I. INPUT VERSUS OUTPUT FILTERS

Allowing social decisions to turn strictly on individual preferences might, in communities of sufficiently bloody-minded individuals, produce some pretty onerous outcomes. Surely it is inadequate to fall back on the purely contingent proposition that vicious preferences are uncommon.⁶ By now, "it would be common ground to nearly all supporters of democracy that there are certain laws or regulations that ought not be passed even if the greater part or indeed the whole of

⁴ In that case, for example, "the concept of a collective choice rule... is itself in doubt" (Sen 1970, p. 85).

⁵ Cf. Sen 1979b, pp. 482–3. Insofar as this extra information about people's utilities itself provides internal grounds for differentially respecting those utilities, this is much more than what Sen (1980–1, pp. 193–4, 207, 210) calls a merely "utility-supported" morality, where the relative weights applied to different components of people's utility are derived from nonutility considerations altogether.

⁶ *Pace Smart* 1973, pp. 67–73.

the people favour them."⁷ In this section, I shall compare alternative methods of imposing and justifying such restrictions on collective choice.

The fundamental contrast is between strategies which filter *outputs* of a social decision function and those which filter *inputs* into it. Whereas output filters work by removing certain options from social consideration, whatever their utility, input filters work by refusing to count certain classes of desires and preferences when aggregating individual utilities. Output filters can be conceptualized as barriers erected at the back end of the social decision machinery, preventing policies predicated on perverse preferences from ever emerging as settled social choices. Input filters might be regarded as barriers erected at the front end of the social decision machinery, preventing perverse preferences from ever entering into consideration.⁸ Or, in more formal terms, output filters act as "stopping modals" telling us "you can't do that"; input filters act to provide "exclusionary reasons" banning considerations of certain sorts of reason for action altogether.⁹

The threat of meddlesome preferences ordinarily drives us to strongly antiutilitarian recourses, such as vested rights. That classically liberal-democratic response is, with the shift "between utility and rights," once again in philosophical favor.¹⁰ The hope is to avoid heinous outcomes by ascribing to each individual a set of rights, thereby circumscribing the application of utility reckoning. Rights function as a "'No Trespassing' sign, . . . a fence erected around an area from which the majority would be excluded by constitutional law"; they create a "protected sphere" and guarantee individuals "protected choices."¹¹

All those phrases seem to suggest that vested rights characteristically act to filter outputs rather than inputs. This emerges especially clearly in the "general theory of rights" offered by Dworkin, who similarly sees the problem with want-regarding moralities as residing in the meddlesomeness of "external preferences":

⁷ Wollheim 1958.

⁸ Input filters also generally come into play earlier, and output filters later, in the social decision process. But an output filter taking effect at the earliest possible moment – excluding certain outcomes from the feasible set, and hence from further consideration, right from the outset – would not be equivalent to an input filter. People's perverse preferences could then still shape their responses to the remaining options, in a way that input filters would preclude; and excluding all options that could conceivably evoke people's perverse preferences prevents them from bringing nonperverse preferences to bear on those options, in a way that input filters would not.

⁹ Anscombe 1978. Raz 1975, chap. 1.2.

¹⁰ Hart [1979] 1983.

¹¹ Mayo 1960, p. 188. Hayek 1960. Hart 1955.

The concept of an individual political right . . . allows us to enjoy the institutions of political democracy, which enforce overall or unrefined utilitarianism, and yet protect the fundamental right of citizens to equal concern and respect by prohibiting decisions that seem, antecedently, likely to have been reached by virtue of the external components of the preference democracy reveals.¹³

On Dworkin's account, rights prevent certain kinds of decisions from emerging *out* of the social calculus, rather than preventing offensive "external preferences" from entering *into* it. And this is, in practice, just how most (if not all) rights work. For the most part, rights restrict results or, at most, procedures. Only very occasionally do they restrict inputs *per se*.¹⁴ Hence, the standard response to the problem of perverse preferences is essentially a variant of the output-filtering strategy.

The first problem with the rights strategy is how to circumscribe and justify the creation of this private preserve. There is, of course, the familiar proposition that the state has no business interfering with "purely private-regarding" actions. But if that means actions affecting no one else, Stephen was right to object that every action has an impact upon someone else besides the actor – there are simply no private-regarding actions in that sense.¹⁵ Nor did Mill claim there were. Rather, he defined private-regarding acts as those not impinging upon anyone else's interests, "or rather certain interests, which, either by express legal provision or tacit understanding, ought to be considered as *rights*." He goes on to say that "the acts of an individual may be hurtful to others, or wanting in due consideration for their welfare, without going to the length of violating any of their constituted rights," in which case the offender may not legitimately be punished by law.¹⁶ Thus, the private-regarding move will not suffice to justify the ascription of rights. Either we take "private-regarding" literally, in which case the private sphere is the empty set; or else we define the private-regarding sphere in terms of rights, which makes it circular

¹³ R. Dworkin 1977, p. 277; cf. R. Dworkin 1978, pp. 134–5.

¹⁴ True, American trial courts suppress evidence obtained in violation of defendants' constitutional rights. But this is not in consequence of their rights *per se* but rather of a particular strategy (the "exclusionary rule") chosen to enforce those constitutional constraints. There is no logical reason why your right against random searches need strictly entail a right to have illegally seized (but undeniably true) evidence excluded at your trial.

¹⁵ Stephen 1874. We could of course give a narrow meaning to "being affected" as, e.g., "having one's life materially impinged upon" (Barry 1965, p. 63). But there is no reason to protect people's material interests to the exclusion of equally important symbolic ones, protecting their self-respect for example (Goodin 1980, chap. 5; 1982b, chap.5).

¹⁶ J. S. Mill [1859] 1975b, chap. 4, p. 92. Gray 1981, pp. 98–101.

for us then to justify rights in terms of the private-regardingness of actions within that sphere.

A more plausible argument for carving out a private sphere builds upon the notion of "respect for persons." If we follow Benn in conceiving of a person as being essentially "a subject with consciousness of himself as agent, one who is capable of having projects, and assessing his achievements in relation to them," then to respect someone as a person we must allow him free rein (within limits, perhaps) to frame and pursue projects as he will.¹⁶ Rights are ascribed to individuals on this reading of the situation, in order to create and protect this private sphere within which people will be free to engage in that distinctively human activity.

This however, raises a second problem for the rights strategy. If we are truly concerned to show people respect, we must not confine ourselves (as the rights strategy does) to prohibiting degrading policy outcomes. We also show people respect or disrespect through our attitudes and motives, even if they do not culminate in actions. Liberals count all preferences without prejudice – humiliating and degrading ones included – in the social decision calculus. Vested rights guarantees, by filtering some nasty options out of the feasible set, might save the state from *doing* anything that shows citizens disrespect. It has, however, already shown them disrespect by *counting* degrading preferences in the first place. Input filters will be required if we are to prevent the sort of humiliation that comes from the social sanctioning of mean motives.¹⁷

Vested rights have been used as a practical illustration of the output-filtering strategy at work. There are few if any mechanisms already in operation approximating the input-filtering ideal. But as a theoretical option this strategy is familiar enough. There have been regular suggestions that we "censor utility functions" in one way or another, ranging from Rousseau's dictum that people's "particular wills" should be excluded in reckoning the "general will," through welfare economists' proposals to weigh "merit goods" more heavily than is justified by the strength of the preferences behind them alone, to proposals from contemporary political theorists to discount preferences which are "external" in form.¹⁸ Thus filtering inputs is a live,

¹⁶ Benn 1971, p. 8; 1988. See also Fried 1978, pp. 28–9.

¹⁷ Goodin 1982b, chap. 5. Output filters might usefully supplement input filters, serving as a second line of defense in case anything slips through the first. For purposes of protecting self-respect, however, they can never substitute fully for them (pace Dworkin 1977, chap. 12; 1978, pp. 134–5).

¹⁸ Harsanyi 1977a, p. 62. Rousseau [1762] 1973, bk. II, chaps. 1 and 3; Riley 1986. Musgrave 1968; Head 1974, chaps. 10 and 11; Walsh 1987; Goodin 1989c. Barry

albeit neglected, theoretical option. And it is one that certainly ought to be pursued, because it is a better response to the concerns that threaten to drive us from preference-based moralities to rights-based ones.

II. GROUNDS FOR LAUNDERING PREFERENCES

The great challenge lies, of course, in finding acceptable grounds to justify the laundering of preferences.¹⁹ Here I shall concentrate on the most unobjectionable form, namely, grounds internal to the preferences themselves. Provided we are prepared to consider richer utility information, much laundering can be justified in strictly want-regarding terms. Taking into account why individuals want what they do, what else they want, and the relationships between and implications of their various desires, five especially interesting justifications for censoring utility functions emerge.

A. Protecting preferences from choices

A person's choices do not always perfectly reflect his preferences, as has long been known.²⁰ As was seen in Chapter 8, people sometimes make their choices on the basis of incomplete information, in ignorance of their own future desires or in the absence of a "full and vivid awareness" of all the alternatives; other times, people's choices merely reflect the framework within which they are set or a desire to avoid responsibility.²¹ In all such cases, we can serve a person's "real" preferences only by censoring the misleading indication of his preferences that is revealed in his choices.²²

1965, pp. 62-6; R. Dworkin 1977, chaps. 9 and 12; 1978, pp. 134-5; 1981, pp. 196-204.

¹⁹ Cf. Arrow 1963, p. 18.

²⁰ See, e.g., Sen 1973; 1977a, pp. 327-9; Gibbard 1986.

²¹ Goodin 1982b, chap. 3. Gibbard 1986. Brandt 1979.

²² Maybe there is something fishy about forcing people to consummate Pareto-optimal deals (Broome 1978a, p. 316; Barry 1986). There is a class of market choices we intuitively feel should be left to individuals to botch as they will. Where we are aggregating preferences into collective choices, however, we are operating well outside that realm. There, it would be even more fishy to feed into collective deliberations misleading preferences revealed by people's suboptimal choices - assuming, of course, some other truer reading of their preferences can be obtained (by directly asking them, or by indirectly inferring them, after the fashion of Tversky and Kahneman [1981]).

B. Reciprocal forbearances

Much of social morality has now been explained in terms of reciprocal forbearances. Many of the things that would benefit one person would harm another person even more; so each agrees to forgo his opportunities for gain at the other's expense, on condition that the other likewise forgo opportunities for gaining at his expense. The much-discussed strategic problems associated with concluding and enforcing such agreements need not detain us here. We need only note that reciprocal forbearances might apply to preferences as well as to actions. People may reciprocally agree to a rule of mutual tolerance, each forgoing his meddlesome preferences on condition that all others do likewise. Wherever people agree (explicitly or implicitly, actually or hypothetically) to such mutual forbearances, laundering their preferences accordingly is fully justified in terms of their own larger preference ordering.

C. Explicit preferences for preferences

On some accounts, people are distinguished from lower forms of life precisely by the fact that they have not only preferences but also preferences for preferences.³³ These come in various forms: as moral principles, personal ideals, social ideals and role responsibilities.³⁴ Such higher-order preferences can justify censoring the utility functions of anyone possessing them whenever, through "weakness of will," those persons fail to live up to their own second-order preferences. If someone finds that one set of preferences is actually guiding his behavior, when he dearly wishes another would instead, then we can justify laundering his preferences as a simple case of respecting his own preferences for preferences.³⁵ In aggregating preferences, we count only those that the individual wishes he had, and we ignore all those he wishes he did not actually experience. No violation of want-regarding principles is entailed, since we choose which preferences to count and which to ignore strictly on the basis of the individual's own (higher-order) preferences.

³³ Frankfurt 1971.

³⁴ Sen 1974a. Goodin 1975. Benn 1976; 1979. Hollis 1977.

³⁵ Rothenberg 1968, p. 330. Of course, if all preferences are on a par and differentiated only according to their strength, then there is no reason systematically to favor someone's weak second-order preference over his strong first-order one (Hollis 1981, p. 176). To justify that we must appeal to notions of people's self-images and self-respect (Goodin 1982b, chap. 5) or a distinction between their "evaluational" and "motivational" structures (Watson 1975; 1977).

D. Implicit preferences for preferences

People only occasionally acknowledge explicitly their preferences for preferences. We can, however, often find logical relations among those preferences that people *do* acknowledge which imply certain other preferences that they may fail to acknowledge. Admit it or not, these are still *their* preferences in some important sense. Just as explicit preferences for preferences can justify ignoring certain first-order preferences in conflict with them, so too might certain logically central implicit preferences override (and justify decision-makers in ignoring) certain explicit but logically more tangential preferences. The social decision-maker would be ignoring someone's first-order preference so as better to serve his own higher-order (albeit implicit) preference. Preferences still dictate decisions – no extraneous considerations are being introduced – so welfarism and want-regarding principles are not violated.²⁶

Many weakly paternalistic arguments appeal in precisely this way to implicit preferences for preferences to justify revising people's statements of their preferences. Gerald Dworkin asks us to "suppose that there are 'goods' such as health which any person would want to have in order to pursue his own good – no matter how that good is conceived. . . . Then one could agree that the attainment of such goods should be prompted even when not recognized to be such, at the moment, by the individuals concerned." The example he offers concerns an automobile driver who, although fully and vividly aware of the risks he runs on the roads, refuses to fasten his seat belt. "Given his life-plans, . . . his interests and commitments already undertaken," Dworkin writes, "I think it is safe to predict that we can find inconsistencies in his calculations at some point"; and, therefore, we can use his deeper (albeit only implicit) preference to justify disregarding his stated preference for not bothering to belt up.²⁷

Another example builds on Hegel's "master-slave paradox," which Elster explicates as follows:

The master is caught in a trap, for he can get no real satisfaction from his power over a being that he treats like a thing. The very concept of unilateral recognition is contradictory, as can be seen by thinking

²⁶ R. Dworkin's (1977, chap. 7) argument for "taking rights seriously" similarly starts with those rights which the Constitution explicitly guarantees, and proceeds to find further rights which although not explicitly recognized are nonetheless logically implicit in those which are. Those implicit rights must not only be considered alongside those which are enumerated but also, being logically more central, must actually *override* the others in cases of conflict.

²⁷ G. Dworkin 1971, pp. 120-1. Rothenberg (1968, p. 330) analyzes the case of the drug addict in precisely the same terms.

through the farcical idea of a nation being diplomatically recognized by one of its own colonies. To the extent that the master treats the slave on a par with cattle, he gets no non-economic satisfaction from his power; to the extent that he treats the slave like a human being, he has no power over him.⁴⁸

From this, we can generalize important conclusions about the whole class of "second-party preferences," defined broadly as preferences about other people's satisfactions whether of a benevolent or malevolent bent. One can find it satisfying to have and act upon such preferences about other people only if they are acknowledged to be people, either by one's own behavior or by some larger system of social values. For the altruist this creates no problems. But for the sadist it does. Both his personal dealings and his first-order preferences tend to degrade and humiliate people, to reduce men to the level of beasts or worse; if he gets his way on policy, however, he will find it no more satisfying to kick a man than a dog, since the social recognition of a man's greater dignity was all that made kicking him more satisfying in the first place. The sadist's preferences for humiliating men in everyday life, then, imply a preference that others' dignity be acknowledged and respected in public policy. Even though he insistently demands the contrary, the social decision-maker is obliged to respect this logically central implication of the sadist's first-order preferences. His utilities are still controlling, however. No violation of want-regarding principles has occurred.

E. Internal logic of preference aggregation

All the previous arguments justify ignoring some of a person's preferences in terms that are somehow internal to his own preference function. A final class of justification points instead to reasons which are located in the logic of the social decision process. Elsewhere I have shown how one such implicit commitment constrains social decisions. Respecting people's choices implicitly commits us to respecting their dignity, and this implies certain clear limitations on the choices we may be obliged to respect.⁴⁹ Similarly, our very choice of aggregating preferences as a way of making social decisions carries consequences for the kind of preferences that we can count. "After all," Harsanyi reflects:

the entire basis for [the social decision-maker] i 's interest in satisfying j 's preference is human sympathy. But human sympathy can hardly impose on i the obligation to respect j 's preferences in cases where the

⁴⁸ Elster 1976b, p. 261.

⁴⁹ Goodin 1982b, chap. 5.

Laundering preferences

latter are in clear conflict with human sympathy. For example, human sympathy can hardly require that *i* should help sadists to cause unnecessary human suffering – even if a very large number of sadists could obtain a very high utility from this activity.”

Just as some aspects of our friends' characters are better ignored, so too are some arguments in our fellow citizens' preference functions better ignored.

Taken together, these five arguments suggest that want-regarding principles, suitably expanded, can afford substantial protection against perverse, meddlesome preferences and the heinous outcomes they might produce. Ultimately we may want to launder preferences more thoroughly than we can find internal justifications (like those discussed) for doing, and we will be forced to fall back upon ideal-regarding principles to do so. But at least I hope to have shown that the demand for decisions based on “nonutility information” – and on vested rights in particular – is premature. There is still much unexploited room for maneuver within the want-regarding framework, provided we take a broad view of what “utility information” can include.

III. THE SELF-LAUNDERING OF PREFERENCES IN COLLECTIVE CHOICE

Formal models of collective choice tend to represent it as some mechanical process of aggregating individual preferences. This badly understates the true complexity of the process. Whereas these models usually take preferences as given, for example, classical theories of democracy have always acknowledged that people can and should reformulate their preferences in response to rational discussions in the course of collective deliberations. Instead of working on some *fixed set* of preferences, the social decision machinery changes them in the process of aggregating them.” Furthermore, and more to my present point, the social decision machine does not necessarily work with each individual's *full set* of preferences. This is partly because people find strategic advantage in suppressing some of their preferences. But

¹⁹ Harsanyi 1977a, p. 62. Notice that no one is excluded from the decision arena: Everyone's preferences count, just not *all* of everyone's preferences necessarily count. Harsanyi's argument here parallels J. S. Mill's (1859/1975b, chap. 5) against allowing anyone to sell himself into slavery. If we respect people's choices because we respect their freedom, then that gives us no grounds for respecting their choice to renounce their freedom.

²⁰ Sunstein 1990, chaps. 1–2; 1991. See similarly J. Cohen 1989, Dryzek 1990, Dunleavy and Ward 1981 and Little 1952, p. 428.

those distortions are ephemeral, shifting with strategic circumstances. There is a deeper dynamic, inherent in the very nature of the collective decision process, which induces people to launder systematically their own preferences, and to express only a small subset of their preferences in the form of political demands. This forms the basis for my discussion here.

Basically, collective decision-making evokes different sorts of preferences, because "an individual's response depends on the institutional environment in which the question is asked." Certain kinds of argument, powerful though they may be in private deliberations, simply cannot be put in a public forum. In countries like the United States, "the market is an institution designed to elicit privately-oriented responses from individuals and to relate those responses to each other. . . . The electoral, legislative and administrative processes together constitute the institution designed to elicit community-oriented responses."³ This latter function is served even more dramatically, in both theory and in practice, by Thai village meetings. Anthropological observers report substantial discrepancies between the way in which villagers say they intend to vote on public works projects ahead of time, and the way they actually do vote in the end. Closer examination of what happens during collective deliberations leading up to these votes reveals that "during the course of the decision period the community-oriented values become increasingly important."⁴

The distinction I am drawing is not between a person's "perceived" interests and some other "truer" interests which he discovers in the course of collective decision-making. Instead, it is between multiple preference orderings actually operative within the individual, and which he applies differently according to the context. As Brandt emphasizes in reply to Arrow, "some choices are motivated by the prospect of enhanced *personal welfare* . . . whereas others" – paying taxes is his example – "are motivated by considerations of *moral principle*."⁵ People can be moved to action by either self-interest or benevolence, by either material or moral sentiments, by "ethical" as well as egoistic preferences, by "social" as distinct from "private" preferences."

³ Maass 1966, pp. 216–7. See similarly Musgrave 1968, p. 160.

⁴ Bilmes 1979, p. 174.

⁵ R. Brandt 1967, p. 27; cf. Arrow [1951] 1963, pp. 82–3.

⁶ See, respectively, Hume 1739; 1777, and Smith 1790; Harsanyi 1955 and Goodin and Roberts 1975; and Plamenatz 1973, pp. 155–68, and Bator 1957. Similar themes are pursued in Self 1975, pp. 134–5, MacRae 1976, pp. 138–45, Sen 1977a, Benn 1979, Margolis 1982 and Mansbridge 1990. Of course, altruistic preferences can be represented as egoistic ones constrained by the exigencies of strategic play in an iterated prisoner's dilemma game (Runciman and Sen 1965, p. 557), or they can be incorporated alongside egoistic ones in a larger metapreference

This multiplicity of preference orderings matters because, in the context of collective decision-making, people will launder their own preferences. They will express only their public-oriented, ethical preferences, while suppressing their private-oriented, egoistic ones. There are various reasons for this. One explanation turns on the logic of "role rationality."⁶ Publicly oriented preferences are somehow uniquely appropriate to the role of the individual *qua* citizen. An individual can hardly conceive of himself as a person who occupies a certain kind of role, and who has the sort of preferences that go with it, without also being disposed to act upon them when required by the role to do so.

Paralleling this argument for moral self-expression couched in terms of roles and self-images is another even stronger argument growing out of the so-called paradox of voting. The paradox is that anyone bothers to vote at all, since the chance of his casting the decisive vote "is less than the chance he will be killed on the way to the polls."⁷ That certainly makes it irrational to vote, insofar as your pay-offs depend upon the outcome of the election. Privately oriented egoistic preferences are of that sort and, therefore, cannot motivate your vote. Ethical preferences, however, are not necessarily of that sort. "Doing the right thing" may bring satisfaction in itself, even if the ultimate outcome is unchanged.⁸ Repudiating some nasty business can enhance your self-respect and sense of integrity, even if it goes ahead without your endorsement.⁹ Why people bother going to the polls at all, and how they vote once they get there, can both be explained in terms of taking a moral stand.¹⁰

The power of a moral crusade to draw votes is well known, certainly among politicians themselves. Among political scientists, there has been some talk of a fundamentally "public-regarding" ethos underlying voting behavior; and it has been shown that "symbolic val-

ordering (Benn 1979, p. 302; Brandt 1967; Sen 1977a). But the former is only part of the story (Goodin 1982b, chap. 6), while the latter representations leave out much that is of significance, behaviorally and otherwise.

⁶ Benn 1979. Hollis 1977. Goodin 1975.

⁷ Skinner 1948, p. 265. Downs 1957, chap. 14.

⁸ Goodin and Roberts 1975. Brennan and Lomasky 1993.

⁹ B. Williams 1973a, pp. 108-18. Hill 1979. Goodin 1982b, chap. 5. Since one vote taken by itself is of virtually no electoral consequence, voters are precluded from acting on consequentialistic principles for consequentialistic reasons. That may lead people to behave more fanatically politically than personally (Hollis 1981). Still, people may embrace consequentialistic ethics as a matter of principle and abide by them even where they carry no (direct, immediate) consequences.

¹⁰ Downs (1957, chap. 14), seeing the paradox, suggested that people go to vote out of a sense of "civic duty." Although Downs was roundly criticized for introducing this *deus ex machina* to save his theory, it can in this way be derived from some other plausible assumptions about the nature of preference functions.

ues" are far more powerful than "self-interest" as a predictor of Americans' presidential votes.⁴¹ Perhaps the most compelling example, however, concerns the tendency of voters to reelect politicians who have presided over periods of economic prosperity. Careful analysis of the data suggests that people are responding not to improvements in their own personal finances but rather, in public-spirited fashion, to improvements purely in the national economic climate.⁴² The explanation I would offer for such findings (though not the one offered by the authors themselves) is that the "paradox of voting" has forced voters to take a morally principled rather than narrowly self-interested stand.

For a practical illustration of citizens using different sorts of preferences in their private and public capacities, consider the problem of time-discounting. One of the most compelling reasons for an individual to discount future costs and benefits relative to present ones is that, come the time, he may not be alive to experience them. How heavily he should discount payoffs depends on how far they are in the future and, of course, on his own life expectancy. The risk of death would, according to midcentury actuarial statistics, imply a discount rate of around 0.4 percent for a forty-year-old American; for an Indian of the same age, it would be 2.15 percent. From a purely individual point of view, it would be foolhardy *not* to discount future payoffs by at least this much. But risk-of-death discount rates have no place in social decision-making. "Mortality probabilities of specific individuals become irrelevant" there, because "the society goes on forever."⁴³

Eckstein, like Pigou before him, thinks this provides the basis for "dictatorial" action on the part of social planners in overriding citizens who try to impose their inappropriately personal discount rates on social undertakings.⁴⁴ In practice, however, this seldom seems necessary. What Pigou and Eckstein overlook is the possibility that "an individual's time-preference map may be strongly influenced by his expected life span in his capacity as an economic man, although in his role as a citizen his time-preference map may reflect great concern for adequate protection for posterity."⁴⁵ When considering social projects, people do seem to take a far longer view than would make sense in their private lives, voting for projects that would hardly begin to pay off until they have long been in their graves. The reason, I suggest,

⁴¹ Banfield and Wilson 1964. Sears, Lau, Tyler and Allen 1980.

⁴² Kinder and Kiewiet 1981; Kiewiet 1983.

⁴³ Eckstein 1961, pp. 457, 459. Given the way people's values and preferences will change over this period, their discount rate should be even higher than this risk-of-death calculation implies (Parfit 1976, pp. 98-9; 1984, appendix F).

⁴⁴ Pigou 1932, pt. 1, chap. 2.

⁴⁵ Marglin 1963, p. 98.

is that in the context of collective decision-making people operate on the basis of social/ethical preference functions rather than upon their private/egoistic ones.

IV. BEYOND SELF-LAUNDERING

The laundering of preferences is easily accomplished – and even more easily justified – when in the context of collective decision-making people do it themselves.⁴⁶ But some perverse preferences (“principled malevolence,” e.g.) would still slip through. Such further laundering as is required must be undertaken by policy-makers, acting on behalf of those with whom they interfere. This paternalism is gentled by two special features. First, overriding a person’s preferences is (on arguments A–D in Section II, at least) justified in terms of that person’s *actual* preferences, rather than in terms of some reading of his “true interests” which he does not himself share. And, second, this argument suggests only that we *disregard* certain votes a person might cast or demands he might make, rather than heavy-handedly force him to do or restrain him from doing something contrary to his will.

The practical political problems surrounding any form of paternalism remain. We need institutions with enough slack between citizen demands and social decisions to allow policy-makers to override mistaken preferences when necessary, but not so much as to cut politicians free from the constraints of citizen preferences altogether. These difficulties are lessened where (as in arguments A–C in Section II) people will eventually come to be grateful to a paternalist who has truly rendered them a service. Then democratic accountability – understood as a post hoc check – can suffice. This allows public officials to serve the “public interest,” understood as what the public will eventually come to approve, rather than cater to the public’s possibly mistaken ideas of what it might like *ex ante*.⁴⁷

Another modest but nonetheless useful method of laundering preferences, which can be used when that one cannot be, is suggested by the analysis in Section I. There I argued that our paramount goal should be to protect people’s self-respect and dignity, and that these are offended by the social sanctioning of mean motives of others that

⁴⁶ It is not just that self-laundering is a choice that people make for themselves. It is important to add that the constraints in response to which they do so are themselves morally acceptable, deriving as they do merely from the need for each to blend his own will with that of others in reaching a collective decision.

⁴⁷ Goodin 1989c. This process is made easier by the tendency for people to vote “retrospectively,” rewarding or punishing incumbents for what they have done in their last period in office, rather than prospectively cuing on campaign promises (Fiorina 1981).

takes place when perverse preferences are allowed to enter the social decision calculus. That is why input filters (such as preference-laundering) rather than mere output filters (such as vested rights) are required. Notice, however that stopping people from having (or even from giving voice to) such preferences is not what is crucial. Rather, it is stopping the social decision-making apparatus from taking official cognizance of them.

The most straightforward way of doing that is through the rules governing debate within the legislature. Already there are certain limitations on what may be said in the chamber. The rules of the British House of Commons, for example, prohibit "treasonable or seditious language or a disrespectful use of Her Majesty's name" and in the American Congress members may not speak "disparagingly about any state of the Union" or subject the president or other members of their own or the other chamber to "personal abuse, innuendo or ridicule."⁴ At present these courtesies are extended only to other members. But there is no reason, in principle, why they should not be extended to citizens and groups of citizens more generally.⁴

The effect of extending parliamentary privileges in this way must not be overstated. Banning racist talk from the floor of the chamber will not prevent measures designed to disadvantage racial minorities from being introduced and passed. At most, my proposal would prevent explicitly racist justifications for them from emerging in the course of parliamentary debates and from entering into the legislative history of the bills eventually enacted. The real motives of their sponsors may be an open secret. Indeed, they may even boast of their racist intentions outside of the chamber itself. The legislature refusing to take official cognizance of perverse (e.g., racist) preferences would, therefore, amount to little more than a polite fiction. That, however, nowise diminishes the value of the gesture in protecting people's self-respect. It is in the nature of dignities and indignities that they are conferred more through symbols and gestures than anything else.⁵

The political realism of this proposal may well be queried. As

⁴ May 1971, p. 415. Cannon 1963, p. 157; U.S. Congress, Senate 1965, Standing Rule 19.4. Other "rules of decorum" restrict the form speeches can take; see Jefferson (1801, sec. 17) and May (1971, chap. 19).

⁵ Gross 1953, pp. 371-2. Note that this does not amount to giving people a right not to be insulted in parliamentary debates, however. No member can waive his privilege not to be insulted on the floor, nor need a member demand his rights in this regard for the Speaker to intervene (U.S. Congress, Senate 1965, Standing Order 19.4; Cannon 1963, p. 159; May 1971, pp. 424-33); indeed, the offended party is obliged to accept an apology when it is proffered, and if he does not he will himself be taken into custody (May 1971, p. 420).

⁶ Goodin 1980, chap. 5; 1982b, chap. 5. Shklar (1979, p. 19) is right. Hypocrisy has its uses.

Dworkin says, it would be futile "simply to instruct legislators, in some constitutional exhortation, to disregard the external preferences of their constituents. Citizens will vote these preferences in electing their representatives, and a legislator who chooses to ignore them will not survive."³¹ But my proposal does not require politicians to ignore perverse preferences of their constituents, but merely asks them not to voice them in the legislative chamber. The votes of representatives (if not their speeches) can still perfectly reflect constituents' desires.³²

In any case, the political feasibility of any scheme for restricting majority rule must ultimately depend, directly or indirectly, on the acquiescence of the majority itself. We speak as if instituting a scheme of "civil rights" will necessarily "remove those decisions from majoritarian political institutions altogether."³³ But constitutions, of course, can be amended or replaced, and rights within them withdrawn. A politician who persists in playing by the established rules when his constituents insistently demand he refrain from doing so will simply be replaced.³⁴ It is therefore surprising and encouraging that the low popular regard for various constitutional rights has translated into so few serious political demands for their repeal.³⁵ The implication seems to be that majorities with undeniably perverse preferences might, under some circumstances, abide by (and perhaps even appreciate) institutions denying them the opportunity to put those preferences into practice.

To some extent, people put up with such restrictions as a matter of "reciprocal forbearances" (of the sort discussed in Section IIB). Having their own preferences denied – whether by rights guarantees restricting outputs, or by preference-laundering rules restricting inputs – is merely the price they pay for institutions that will similarly deny the preferences of others.³⁶ There is no more point in punishing their representatives for agreeing to those restrictions than there is in punishing them for agreeing to a national budget that devotes less than 100

³¹ R. Dworkin 1978, p. 134.

³² That may not be enough to satisfy their constituents. Myrdal (1944, pp. 60–1) reports that, in the "rank order of discrimination" in the American South, symbolic gestures were always much more important than material interests.

³³ R. Dworkin 1978, p. 134.

³⁴ Goodin 1975; 1992b, chap. 6. Barry 1975, p. 409.

³⁵ Prothro and Grigg 1960. McClosky 1964. McClosky and Brill 1983. Sniderman, Tetlock, Glaser, Green and Hout 1989.

³⁶ The need for reciprocal forbearance might arise from standard prisoner's dilemma-style collective action problems (Goodin 1976; Hardin 1982). Or people might realize that pursuing "status goods" (Hirsch 1976) is self-defeating, and rule out such preferences from the outset. Or they may realize that things like racial hatred are self-fueling (Lave and March 1975, pp. 396–402) and agree that racist preferences be ignored in order to nip that process in the bud.

percent of public expenditures to their own constituency. But there are some "persistent minorities" whose reciprocal forbearances we can be quite sure we will never require, and at the end of the day tolerating rules protecting them can only be explained in terms of internalized moral principles of the sort discussed in Section III.¹⁷

V. CONCLUSION

Social choice theorists, in order to make their subject mathematically tractable, have tended to narrow our understanding of "utility" to a simple, single preference-ordering. Much is thereby gained in terms of formal rigor. But much is also lost. A broader understanding of what "utility information" might consist in would, as I have shown here, serve utilitarians far better. It not only would enable them to make sense of how individuals actually do launder the preferences they express in collective decision contexts but also would justify them in instructing social authorities to respond selectively to only certain sorts of citizen preferences.¹⁸

¹⁷ Goodin 1992b, esp. chaps. 2, 4, 5, 7.

¹⁸ Earlier versions were presented at Brussels and Ustaoset. I am particularly grateful for comments, then and later, from Jon Elster, Dagfinn Føllesdal, Aanund Hyland, Herman van Gunsteren, Francis Sejersted, Cass Sunstein and Mike Taylor.

Heroic measures and false hopes

"Heroic measures" refers to the deployment of unusual (rare, experimental, expensive, nonstandard) technologies or treatment regimes, or of ordinary ones beyond their usual limits. The examples ordinarily offered concern care for the terminally ill, a heart-lung machine hooked up to someone who is brain dead, and so on. But, philosophically, special complications are posed by people who would otherwise (and who may, anyway) cease to be. Here I shall focus instead on cases where such complications are absent – on extraordinary measures for creating life (in vitro fertilization and such like), and on extraordinary measures for improving the quality of the lives of the handicapped (e.g., electrical stimulation of paraplegics' muscles to simulate walking).¹

Such heroic measures represent an unattractive aspect of utilitarianism. They amount, if not quite to killing people with kindness, at least to torturing people with kindness. Yet insofar as it is a kindness, insofar as people's well-being is indeed promoted by the intervention, there would seem to be no utilitarian grounds for ending the torture. This chapter is aimed at correcting that shortcoming. Here I shall be offering considerations, internal to utilitarianism, which militate against deployment of heroic measures. Those considerations do not always prove conclusive. But they must always be weighed in the balance, and they may occasionally (in certain classes of cases, they may even characteristically) tip that balance.

¹ As reported in *Time*, December 13, 1982. The conclusions of this chapter may be applied, *mutatis mutandis*, to the dying – although exactly what *mutatis mutandis* might there involve may prove to be an enormously difficult question.

I

It seems distinctly odd to be arguing against heroic measures. Heroism is something we ordinarily regard as *exceptionally good*. But the term "exceptionally" good is telling. From the point of view of the person performing the acts of heroism, such acts are exceptional in the sense of being above and beyond the call of duty – where, as on Urmson's analysis, the limits of duty are set at what can reasonably be expected of competent moral agents.² Correlatively, from the point of view of the beneficiary of heroic interventions, such acts are also exceptional in the sense of being rare events, since "what can reasonably be expected" evokes (for moralists, as for lawyers) a sense of the "normal" that has a certain frequentist tinge about it.

The sense of the heroic that is carried over from this standard analysis of heroism into the analysis of heroic therapeutic interventions retains an air of the extraordinary and the exceptional. But in that application, the notion of the heroic is broadened somewhat.

There are, in fact, two senses in which medical treatment might qualify as "heroic measures." The first is the same sense as that in which heroes themselves are rare; that is, the requisite performances are made only occasionally. A second sense in which medical treatments might qualify as "heroic" is that it is only the exceptional patient who will, if given the treatment, actually derive any benefit from it. One in a hundred, or one in a thousand patients might benefit from some experimental cancer cure if it were administered equally to all of them, suppose. That, too, would count as a "heroic" form of treatment.

Whichever the sense in which the treatment is heroic or exceptional, however, it is always presumed to be exceptionally good. It is that presumption of goodness which I next turn to query.

II

Heroic measures are presumed always to be good: certainly, at least from the point of view of the recipient; and by virtue of that, at least presumptively from the point of view of society at large. Think of the case of the terminally ill. Without the heroic treatment, they would cease to live. If the treatment in view promises (with *whatever* probability, however small) a life that is worth living for them, then it is at least from their point of view better that they should receive the treatment than not. Any chance of a good life is better than no chance of a good life.

² Urmson 1958.

There is nothing special about the terminally ill in this regard. The same easy presumption of the desirability of heroic measures can be grounded on the same sort of logic for those who are not terminally ill. Suppose the paraplegic's life is worth living; it is not as rich or varied as that of the able-bodied, perhaps, but it is better than nothing. Still, a life with some mobility is, presumably, better than a life spent entirely in a wheelchair; and any chance of a better life is better than no chance of a better life. So heroic measures to enable the cripple to "walk" are presumably good, just so long as the treatment does not entail any downside risk of something even worse than the cripple's present paraplegia (e.g., quadraplegia).

In practice, of course, heroic measures sometimes do offer precisely that: a lottery that has, as its possible outcomes, either a life that is very much better than the one presently being enjoyed or one that is very much worse.³ In cases such as that, heroic measures are neither presumptively good nor presumptively bad. To decide whether they are good or bad, in any particular case of this sort, we must consider the relevant probabilities, the values attached to each possible outcome, and the value or disvalue associated with risk-taking per se.

Those, however, are considerations that I would like to bracket out for purposes of this discussion. For present purposes, let us suppose that heroic measures pose no downside risk whatsoever. In what follows, I shall be supposing that heroic measures will either make the recipients of such treatments better off or else they will leave them no worse off than they would have been in the absence of the treatment.

III

Even after all that has been bracketed out, there still is, I shall argue, a sense in which heroic measures in medicine may leave their recipients worse off. This is through engendering "false hopes." It may be good (albeit above and beyond the call of duty) for us to be heroes. But it is wrong for others to expect heroic performances from us; and it is wrong for us to lead people on in this respect, causing them to expect more heroism than is actually afoot. If the expectation follows inevitably from the performance, then that is an argument against the performance itself. Such is the structure of my argument here.

Heroic measures have been described here as "unusual." They are heroic either because they are not often undertaken, or else because they do not often work to produce the desired effects. In medical applications, especially, these two aspects of heroism are often con-

³ In discussions of "death with dignity" perhaps this is precisely what is at issue – living like *that* is worse than not living at all.

nected. Measures that usually do no good are not usually undertaken, precisely because usually they would not work anyway.⁴ Certainly not all cases fit that mold, and I shall return later to discuss cases where failure to undertake heroic measures cannot be thus justified. For now, however, let us focus upon cases where that analysis is accurate. For those cases, at least, we may legitimately equate "heroic measures" with "low-probability-of-success treatments."

In analyzing what might be wrong about offering people low-probability treatments, it is important to recall that people are *offered* such treatments. People are not forced to accept such treatments; they are not given the treatments unless they or someone acting as their agent consents to it. Now, we regularly let people take risks. Some of them carry a very real possibility of undesirable outcomes; some of those undesirable outcomes entail serious damage to their health or other vital interests. Still, we are perfectly comfortable in letting people engage in those risky ventures, just so long as we are sure that certain basic preconditions are met (their consent is present, it is fully voluntary, well informed, reflects their settled preferences, etc.). Why then should we have any qualms about letting people take risks (even very bad - i.e., low-probability - ones) of *improving* their health, especially where *ex hypothesi* the treatment could *only* improve their health and never worsen it? That, on its face, seems to present something of a puzzle.

My resolution of that puzzle will focus on the role that offering people low-probability treatments plays in engendering in them (or those who care about them) false hopes, and on the way that such false hopes in turn undermine people's welfare.

IV

Before turning to my own argument as to why this is true, let me distinguish mine from a more familiar method of arguing for similar conclusions. It is well known that people's probability judgments are subject to all sorts of unwarranted influences. Building on those familiar propositions, it might be thought that what is wrong with heroic medical interventions is that they lead people to imagine that the probability of success is far higher than it actually is.

The phenomenon here in view amounts to "wishful thinking." People desperately in need of treatment, when no treatment looks particularly promising, resort to long shots. Very much hoping, against all the odds, that the treatment will work, they come to believe that it

⁴ That is the characterization of "standard practice" underlying definitions of malpractice and medical negligence (Fletcher 1968).

will. The wish, rather than the evidence, is what here gives rise to the belief.⁶

On that reading, the problem with false hopes engendered by heroic medical treatments lies in the false beliefs people have about their probability of success.⁶ False beliefs usually carry harmful consequences, at least on average and over the long haul. In the special circumstances here in view, however, they do not. *Ex hypothesi*, the treatment can only (1) make people better off than they would otherwise have been or else (2) leave them precisely the same as they would otherwise have been.⁷

Given that, the long-shot treatment is a "strictly dominant" strategy. It is no worse than any other strategy in any possible state of the world, and better than any other strategy in at least some possible states of the world. Such a strictly dominant strategy, where available, is uniquely rational. It does not matter how dominant it is – how many states of the world there are in which it is actually better than (rather than merely "at least as good as") other strategies, or how likely it is that those preferred states of the world will occur. And since probabilities are of no consequence to strictly dominant strategies, neither are (false) beliefs about probabilities. In short, the penalty that we ordinarily pay for acting in reliance on false beliefs about probabilities is waived, under the special circumstances here in view.

V

In the previous argument, harm was done to people by their beliefs that the probability of heroic measures succeeding was higher than it actually was. The fatal flaw in that argument was that that turns out not to be a harm at all, because in the circumstances here postulated people's choices are unaffected by their probability judgments, distorted or otherwise. The harm that I shall next be considering derives simply from holding out any hope at all to desperate people. When we propose to someone some heroic measure or another to alleviate the problem, we are thereby holding out some hope – some possibility

⁶ B. Williams 1973b, pp. 136–51. Pears 1984.

⁶ Either they do not believe that the probabilities of success are as low as they really are, or else they cannot bring themselves to act upon that belief. One reason may have to do with what psychologists call the "availability" heuristic – the occasional miraculous cure is more memorable, and psychologically registered as more probable, than the multitude of mundane failures (Tversky and Kahneman 1981; Kahneman, Slovic and Tversky 1982).

⁷ Of course, "otherwise have been" embraces the alternative treatments people might have undertaken as well as the course the untreated disease would otherwise have taken. It would be obviously wrong for doctors to offer a low-probability treatment, where some higher-probability one was available.

- of success. That in itself harms people I shall now argue. The harm that comes to people, on this argument, derives from the distortion in their life plans that is produced by the introduction of successful treatment even as a mere possibility. That, rather than any exaggerated notions they may harbor as to its chances of success, is the real source of their injury.

What is wrong with false hopes is that they lead people to pursue illusory goals. Those goals are illusory, first, because they are (probably) unattainable. That, in turn, makes them illusory in a second sense as well. The very goodness of the goal is itself an illusion; and that is true for reasons connected, somehow, to the unattainability of the goal. It turns out to be not merely foolish but positively harmful, in some way or another, to pursue goals that are unattainable.

There are basically two ways of going about analyzing what is wrong with (probably) unattainable goals. One appeals to standards that are external to the agent himself; the other appeals to standards that are somehow internal to the agent himself. Appeal to external standards would be more powerful, if only those standards could somehow be validated. But that, of course, is a tricky business. Internal standards are less contentious and, motivationally at least, no less compelling. The bulk of my argument will therefore be couched in those terms, after only a brief nod in the direction of the sorts of external standards which might be employed in this connection.

The external standard to which arguments against false hopes might appeal is just a variation on the classic stoical argument that people should (for the sake of their own happiness, or peace of mind) revise their desires in light of what they can realistically expect to get. This principle requires only a little revision in the present context. The point of the stoics is couched in terms of possibilities. People ought not to desire that which is impossible. What is at issue with heroic therapeutic interventions is improbable, not impossible, goals. But the basic stoic point retains its force there, too. People will only make themselves miserable pursuing goals that are probably beyond their reach, or devoting more effort (or attaching greater hopes) to goals than their objective probabilities of attainment truly warrant.

Whether or not we can appeal to external standards for these purposes, it is clear that we can make a very powerful appeal indeed to internal standards. And this strong sort of internal appeal must be distinguished from the relatively weak form of appeal that addresses only those who explicitly want to be realistic from the start. Some people do, others do not. Any argument that is hinged on brute facts about people's first-order preferences in this way would be powerfully compromised by their variable first-order preferences in this matter.

The more powerful form of internal appeal turns instead upon reasons people have for regarding heroic measures as *prima facie* undesirable, whether or not they actually do. These are reasons that are still, in some sense, internal to people's own existing preferences. But they are reasons that lie beneath the surface level of their first-order preference themselves.⁸

The key to this argument is just this. Life plans are complex things with various interactions between the particular projects that comprise them. Heroic interventions, if successful, would make a big difference to those plans.⁹ Their success would require a comprehensive revision of those plans, not just some marginal adjustments. If a cripple could walk, his life would change completely; if a childless couple had children, their lives would change completely; and so on.

Thus, people with the prospect of heroic interventions find themselves at a fork in their lives. Depending on how things turn out, they will want to pursue very different life plans. These are typically incompatible alternatives. They must pursue either one or the other, and cannot hedge their bets by pursuing a little of both. And typically choices between such alternatives are substantially irreversible. Once having set out down one path, it is costly if not impossible to go back to the fork again and start out down the other instead.¹⁰ In any case, even where hedging and backtracking might be possible, there is a good reason to avoid it. A life of equivocation and false starts is a less good life than is one characterized by more coherence and consistency.¹¹

The upshot is that people facing the prospect of a heroic intervention must hold all their other plans in abeyance pending the outcome of those interventions. The final step in the argument against heroic measures is to say that that waiting is itself costly to people. It may entail actual out-of-pocket expenses. But at the very least, it will entail "opportunity costs." There are various other projects that people could be pursuing, but are not pursuing, pending the outcome of the heroic intervention.

Of course, the price of waiting is a price that people would only

⁸ The argument sketched below is superior to another, dealing simply in terms of "sour grapes" – not wanting something if you (probably) cannot have it. That is as irrational in just the same way as wishful thinking (Elster 1983, chap. 3).

⁹ We would not call them "heroic" if they did not. Semantics aside, we would have no grounds for taking an action with such low chances of success unless the change it might make would be substantial.

¹⁰ Arrow and Fisher 1974. Of course, heroic measures themselves can sometimes be option-preserving strategies.

¹¹ Nozick 1981, pp. 403–50. Wollheim 1984.

too happily pay if the interventions turn out ultimately to succeed. My point is just that it is a price that people have to pay for heroic interventions, whether they succeed or fail.

The structure of my argument in Section IV, in replying to the probabilistic form of criticism of heroic interventions, was to say that heroic interventions as here construed constitute no-lose propositions. Either they leave you better off than you would otherwise have been or else they leave you exactly as you would otherwise have been.

The structure of my argument in this section is to say that, because that lottery is not played out instantaneously, heroic interventions always constitute costly propositions. Succeed or fail, they always entail interruption to life choices pending their outcomes. How great those costs might be varies.¹⁴ Whether those costs are worth paying varies. It is no part of my claim that heroic medical interventions are always, on net, disadvantageous. My claim is merely that there are always these costs to be weighed in the balance.

(As I said at the outset of this chapter, the terminally ill constitute a special case. Now we can see why. According to the conventional wisdom, anyway, the terminally ill can have no alternative plans to hold in abeyance pending the outcome of heroic interventions, and in this way are radically unlike others who might be offered such interventions. If the intervention fails the terminally ill die, whereas if *in vitro* fertilization fails the couple carries on their very different life without children. Of course, this conventional wisdom misconstrues the situation of the terminally ill. They will typically want to make plans – euphemistically, to “arrange their affairs” – even if it is only to die well. And in any case, the life choices of families and friends will be held in abeyance pending the outcome of a terminally ill patient’s treatment, even if he himself has no plans that are being held in abeyance. More generally, taking the terminally ill thus misconstrued as the paradigm case for heroic interventions quite wrongly implies that those offered heroic medical interventions typically have no choice – or at least no acceptable choice – but to hope that the long shot succeeds. Even as regards the terminally ill that is not true, as advocates of hospices and holistic medicine rightly argue against advocates of aggressive surgery. It is all the more untrue as regards childless couples or paraplegics, who really do have possibilities for perfectly reasonable lives even without heroic interventions.)

The argument developed so far constitutes a case for an individual’s rejecting heroic measures in his own case, on the grounds that

¹⁴ With, among other things, the time it takes for those uncertainties to play themselves out – if there is always “one more possibility” just around the corner, then people’s life choices may be suspended virtually indefinitely; if there are few possibilities and quick resolutions, the interference might be slight.

they entail costs in excess of their probable benefits.¹¹ But we can build on that proposition to come to a case for rejecting certain whole classes of heroic measures quite generally, as a matter of public policy rather than merely personal choice.

The bridging proposition required here is much like the psychological proposition developed in Section IV. Wishful thinking being what it is, people will, if offered the prospect of a heroic medical treatment, imagine that the treatment has a better chance of success than it really does. In consequence, they will weigh the probable benefits of the treatment too heavily, relative to its costs; and they will make what, by their own standards of value, is the objectively wrong choice. In Section IV, their misestimates of probabilities did not matter, because heroic measures were seen as costless. Having here come to see how errors of this sort can be costly, making mistakes about probabilities can indeed cause harm. Here, there is a very real penalty to be paid from making the wrong choice, and that is what mistaken probability assessments trick people into doing.

Roughly speaking, the broad classes of heroic medical interventions that might be less attractive in that light would be ones displaying either (or both) of the following characteristics. One is that the ultimate outcome of the treatment regime will not become apparent for quite some time. The other is that those who would receive the treatments would have reasonably good alternative life choices available to them, even in the absence of the treatment. The former consideration speaks to questions of costs of holding other plans in abeyance for a very long time while the consequences of the treatment play themselves out; the latter speaks to questions of benefits of the treatment regime, compared to those that would be realized even if that treatment were withheld. The higher the costs or the lower the benefits of a treatment regime characteristically are, the more wary we ought to be of offering it to people at all.

VI

None of this constitutes an ironclad argument against heroic medical interventions, either in general or as applied to particular treatments or particular patients. At most, I can only claim to have shown that there is always a price to be paid for such interventions, and that there ought therefore to be a presumption against rather than in favor of such interventions. But this presumption, like all presumptions, is em-

¹¹ More formally, heroic measures should be rejected if the opportunity costs of waiting to see how they turn out would exceed the utility of success, discounted by its improbability of succeeding, less the utility of the best alternative life plan absent the heroic measure.

inently rebuttable. Any given person may suppose the price worth paying. And say what we will about the general propensity among people toward wishful thinking, any given individual may be seeing matters perfectly clearly in so judging.

All that to one side, however, there is yet another way in which heroic medical interventions might be justified, notwithstanding all the arguments lodged against them here. Suppose that heroic measures involve treatments with a low probability of success, as suggested previously. But further suppose that the probabilities of success improve with each successive trial; that is, doctors learn from their experience. And suppose, finally, that that is the only way that the doctors can learn what they need to know in order to perfect the treatment.

The heroic medical interventions might be justified in roughly the same way that we justify using the terminally ill as subjects for trials of long-shot cancer cures. Someone has to be the guinea pig. Giving these patients the treatment engenders in them false hopes; but someone has to be given false hopes in order that subsequent patients can have a real hope.

We are naturally – and quite rightly – queasy about using the terminally ill in these ways, necessary though that may be. We build in various safeguards, and impose various conditions, before this can be done to them. Surely we ought to hedge heroic medical interventions, thus rationalized, with similar constraints. Perhaps the most central of these, for present purposes, is the requirement that we have *some* reason (analogous to promising results from animal experiments, in the case of cancer cures) for supposing that the heroic intervention might do some good, even for the present patients who are basically acting as guinea pigs. That is just to say, it may be an unfortunate necessity to engender in people largely false hopes, but it is never permissible to engender in them hopes that are *entirely* false.

Notice, finally, that it is the *promise* of treatment, rather than treatment *per se*, that generates false hopes. The paraplegic hopes that technology will soon develop to enable him to walk. The *in vitro* fertilization patient hopes that it works next time. And so on.

Some heroic medical interventions entail such promises for the future. Others do not. Imagine the case of a stagnant technology, with no prospect in either the short or even medium term of any breakthroughs. Suppose further that a single application of this “heroic” technology will decisively determine the results that any given patient might expect from it; if it does not work the first time, there is no point trying it another time. Suppose further that all patients who could possibly benefit from such treatment are given it immediately on being identified, so there is no waiting involved. *Then* application

of these "heroic" technologies would resolve uncertainties rather than engendering them, and would be immune to the criticisms lodged against them in Section V.

Needless to say, most technologies are not like that. In the real world there are – and probably always will be – queues. With most treatments, no single trial is conclusive for any given case – if only because the medics might have misapplied the technology or misread the results. And most technologies develop in unpredictable ways at unpredictable rates – we can never be sure that the technology is stagnant, and useless for a person in the future if it is proved to be useless for him now.

With respect to queues and mistakes, we might be inclined to express the pious hope that they be eliminated. If they are, those objections to this strategy for rescuing heroic medical interventions from the criticisms here lodged against them would drop away. But no one can wish an end to breakthroughs in technologies that enhance the quality of people's lives. So even in the real world, some of my criticisms of heroic medical interventions for raising false hopes will retain their force.

Again, nothing I have said should be taken as a conclusive argument against medical innovation. My argument is merely that there is always a price to be paid, in terms of the false hopes it engenders and the harm that they do to people. But that may be a price that we should be prepared to pay.⁴

⁴ I am grateful to audiences at the Universities of Stockholm and York for comments on earlier versions of this chapter.

Theories of compensation

From a moral point of view, the function of compensation is straightforward. Compensation serves to right what would otherwise count as wrongful injuries to persons or their property. That is the role of "compensatory damages" in the law of torts and of "just compensation" for public takings of private property. And that is what the New Welfare Economists were relying on when making the possibility of gainers compensating losers the proper measure of permissible policies.¹

It would, however, be wrong to presume that we as a society can do anything we like to people, just so long as we compensate them for their losses.² The subset of the policy universe to which such a proposition properly applies – policies that are "permissible, but only with compensation" – is bounded on the one side by policies that are "permissible, even without compensation," and on the other side by policies that are "impermissible, even with compensation."³

There clearly are some things that we as a society can do to people without compensating them in any way for their ensuing losses. The state need not compensate people who are stopped from endangering the public health, safety or welfare.⁴ No one expects state inspectors to compensate owners of unsanitary restaurants or unsafe factories

¹ Prosser and Wade 1979, sec. 903ff. Michelman 1967; Ackerman 1977. Kaldor 1939; Hicks 1939.

² Or, in the hypothetical formulation of the Kaldor-Hicks principle, could compensate them for their losses.

³ My focus here is on what public officials may legitimately do to individuals. Analogous issues arise in deciding what individuals may legitimately do to other individuals (Nozick 1974, p. 59).

⁴ This emerges in American constitutional law as the distinction between actions arising under the state's "police power" and its "takings power." See Michelman 1967; Ackerman 1977; Sax 1971 and, more generally, Corwin 1978 and Tribe 1978, pp. 461ff.; cf. Epstein 1985.

they close down. No one supposes that the legitimacy of public health authorities putting victims of smallpox into quarantine is in any way contingent upon compensation being paid to them for lost wages. No one expects the police or courts to compensate the murders or thieves they incarcerate.⁵ No one expects the legislature to compensate tax accountants when passing new legislation to close a lucrative loophole in the present code, or owners of gas-guzzling cars when increasing the gasoline tax, or taxpayers generally when levying a new tax.⁶ Such actions as these, taken under the state's police or taxing powers, are perfectly permissible, even without compensation being paid to those who lose as a result.⁷

The converse is also true. There are some things that we as a society cannot do to people, even if they are compensated for their resulting losses. This class of cases provides the principal focus for the present chapter.

When trying to carve out a case for absolute prohibitions, earlier writers have usually tended to argue that some policies are impermissible because it would be impossible to compensate people fully for their resulting losses.⁸ One tack is to say that the losses would be

⁵ True, those who are quarantined or imprisoned are ordinarily paid a small *per diem*, collectable upon discharge. But this modest sum rarely constitutes anything approaching full compensation, even just for earnings (even just legal ones!) lost while detained. As evidence, witness the way in which those successfully suing for false imprisonment get far more, even in purely "compensatory damages," than the *per diem* due anyone imprisoned, whether rightly or wrongly.

⁶ At least not usually. For rationales and qualifications, see Epstein 1985, chap. 18; Feldstein 1976b; and Sidgwick 1891, p. 188. Neither do owners of a building, such as Grand Central Station, need to be compensated when its being declared a Historical Landmark precludes them from building an office block atop it. Nor do we expect people to be compensated, by one another or the public at large, for losses inflicted by ordinary operations of economic markets (Haveman, Halberstadt and Burkhauser 1984, p. 32; cf. Blough 1941).

⁷ How to distinguish these two classes of cases lies beyond the scope of this chapter. Although it may be true no compensation is ever due to violators of rights (Sidgwick 1891, p. 187), it is not true that compensation is due only to people whose rights have been violated – compensation is due in cases of voluntary sale, for example, even though no rights were violated (Thomson 1986, p. 77). Allied to that is an explanation in terms of "legitimate expectations." We need not compensate people when depriving them of things that they had no reason to expect they would be able to keep; we do need to compensate them when depriving them of things that they had no reason to expect would be taken away. Yet another analysis is that those public activities not requiring explicit compensation are ones forming part of a larger social contract from which everyone in society derives "implicit in-kind compensation"; those for which explicit compensation is due are those of which that is not true (Epstein 1985, chap. 14).

⁸ This is the approach of Nozick (1974, pp. 66ff.; cf. Goodin 1976, p. 81). Alternatively, we might try to ground a case for prohibitions on considerations of effi-

infinite, and impossible to compensate for that reason; another is to regard the losses and compensations as incommensurable, so we could never know whether compensations were adequate to cover losses.⁹

Both those approaches have run into serious difficulties. As a purely practical matter, it is difficult to adjudicate conflicts between two things, each of which has infinite value.¹⁰ More fundamentally, infinite values imply lexicographical priority rules, which are wildly implausible. There are few (if any) pairs of goods such that we would refuse to sacrifice any quantity, however small, of the more valuable to secure any gain, however large, in the less valuable.¹¹

The "incommensurability" approach misrepresents hard choices as easy ones. We may find it hard to say whether Sartre's student should abandon his aged mother to fight with the Resistance or abandon his country to stay and comfort his mother; but whatever else we say about this choice, we are confident that it is rightly to be regarded as a *hard* choice. Representing the competing claims of kin and country as incommensurable would carry the opposite implication. Since no one solution would then be demonstrably any better than any other, the student might as well just flip a coin rather than agonizing over the choice. That, however, is surely *too easy*.¹² Wherever we are tempted to say that the values at stake in some choice are incommensurable, we are likely to be similarly uncomfortable with such a trivialization of a choice that we think should rightly be regarded as tragic.

Here I shall take a different tack altogether. I shall not be saying that policies are impermissible because compensation is impossible in either of these ways. I shall concede that compensation in some sense can be paid. But that is compensation in a *different* sense than that which renders permissible otherwise impermissible policies. For that transformation, compensation of a *strong* sort is required. In the class of cases here in view, only compensation of a *different* and much weaker sort is available.

ciency or distributive justice, rather than on the impossibility of compensation (Calabresi and Melamed 1972).

⁹ Williston 1932, sec. 361. Zeckhauser and Shaefer 1968, pp. 38ff. Tribe 1972, pp. 87ff. Feinberg 1973, p. 92. B. Williams 1981, chap. 5.

¹⁰ Unless one wants to talk in terms of "different-sized infinities" (Feinberg 1973, p. 92 n. 8; cf. Mishan [1970] 1974, p. 462).

¹¹ Sen 1974b. Harsanyi 1975. See more generally Nozick 1968.

¹² C. Taylor 1976, pp. 290-1; see further Griffin 1977; B. Williams 1981, pp. 76ff. Similar objections are lodged against the "anything goes" implication that seems to follow from Kuhn's arguments about the incommensurability of scientific paradigms (Lakatos and Musgrave 1970).

Of course, some other right-making characteristic might always intervene to render a policy permissible even if the right form of compensation is unavailable to do so. It is no part of my thesis that all policies not admitting of this strong form of compensation are necessarily illegitimate *tout court*. My thesis is merely that arguments couched in terms of compensation cannot, in these cases, provide the needed legitimization.

I. THE NOTION OF "COMPENSATION"

A. Compensation in general

The general idea of "compensation" is straightforward enough. To compensate someone for something is, in the words of the U.S. Supreme Court, to provide that person with "a full and perfect equivalent" for that thing.¹¹ If he is given more than that, we would say he has been "overcompensated"; if less, "undercompensated." Being bracketed as it is between these other two notions, the notion of compensation *per se* clearly implies the providing of the *exact* equivalent – neither more nor less.

To compensate someone is to provide him with something that is good, that is, with things that are desired (or at least are desirable).¹⁴ The aim is to bring him up to some baseline of well-being. That baseline to be used for reckoning the adequacy of compensation will typically be identified by reference to some status *quo ante*, that is, some position that the individual himself actually enjoyed at some previous time. Thus, in the law of torts, the baseline for compensatory damage calculations is the position that the injured party was in before the tort was committed against him; when property is taken under the government's power of eminent domain, the compensation due is reckoned as the amount of the property-owner's loss, understood as the difference between his position in the baseline situation prior to the seizure and his position afterward; and so on.¹⁵

¹¹ Brewer 1893, p. 326. See also Atiyah 1980, p. 5; Atiyah and Cane 1987, p. 5; and Day 1981.

¹⁴ Inflicting harm or revenge, although sometimes characterized as "negative compensation" (MacCormack 1973), cannot count as compensation in this sense – unless, of course, the person being "compensated" is made better off by that other person's being made worse off, as sometimes happens (Hirsch 1976; Sen 1983).

¹⁵ Prosser and Wade 1979, secs. 903ff. Atiyah 1980, chap. 7; Atiyah and Cane 1987, chap. 7. Corwin 1978, p. 402. Occasionally the baseline is some independent norm or ideal which, although perhaps standard among some of the population as a whole, was never previously enjoyed by the individual being compensated. It is only in this attenuated sense that we "compensate" the congenitally handicapped for vision they never had by providing seeing-eye dogs, or the educa-

Finally, notice one further general point. Compensation is not the same as restitution. It is one thing to restore the object itself to its proper owner. That is what we (and the *Oxford English Dictionary*) call "restitution." It is quite another thing to compensate the person for its loss (much less for its absence all along). Such compensation is characteristically a matter of providing something which will, in the words of the *Oxford English Dictionary*, "counterbalance, neutralize, or offset" the loss or disadvantage. What all those terms suggest, in turn, is not the restoration of the object itself, but rather the provision of something else altogether.¹⁶

B. Two kinds of equivalence

The central claim of this chapter is that there are two kinds of compensation. These correspond to the two fundamentally different ways in which one object can constitute an "equivalent" for another object which the person has lost (or never had).

The first kind of compensation might be called *means-replacing compensation*. The idea here is to provide people with equivalent means for pursuing the *same* ends (the same as before they suffered the loss, or as they would have pursued had they not suffered the disadvantage). Giving someone who has been blinded a sighted amanuensis or someone who has lost a leg an artificial limb are attempts at this kind of compensation, which I shall hereafter call compensation.

The second kind of compensation might be called *ends-displacing compensation*. The idea here is to compensate people, not by helping them pursue the same ends in some other ways, but rather by helping them to pursue some other ends in a way that leaves them subjectively as well off overall as they would have been had they not suffered the loss or disadvantage at all. Giving someone who has suffered a bereavement an all-expenses-paid Mediterranean cruise might be an example of this sort of compensation, which I shall hereafter call compensation.

The distinction between these two kinds of compensation might be summarized thusly. The first kind of compensation attempts to provide people with equivalent means to the same ends. The second kind of compensation attempts to provide them with equivalent satisfactions through different ends.¹⁷

tionally disadvantaged for stimuli that they never had at home by providing preschool education (Culyer 1974, pp. 22-3; Haveman et al. 1984, p. 30).

¹⁶ Sidgwick 1891, p. 180. Similarly in cases of congenital disadvantage, corrective lenses compensate for ill-shaped eyeballs, hearing aids for weak eardrums, just as cash payments constitute "just compensation" for confiscated lands.

¹⁷ This parallels Atiyah's (1980, pp. 537-8; Atiyah and Cane 1987, pp. 474-6) dis-

Both standards of compensation insist that people must be made *as well off* as they would have been, had it not been for the loss or the disability for which they are being compensated. With compensation, however, they will be made as well off but *differently off* than they would have been. To achieve compensation, it is not enough that they somehow or another be made as well off. They must be left *identically* situated with respect to exactly the same sets of ends.

II. COMPENSATION IN PRACTICE

In due course I shall argue for the moral superiority of compensation over compensation. In attempting to motivate that argument, however, it might be useful first to reflect upon compensation as it is currently practiced in public policy. Contemporary societies have developed a wide variety of ways for compensating people for all manner of accidents, injuries, illnesses, disabilities, losses, and so on. In surveying them all, it is striking how many of our public policies aim at what I have here called compensation, and how few aim at compensation.

The distinction is never phrased in precisely those terms, of course. Instead, lawyers typically distinguish between compensation for pecuniary harms and for nonpecuniary ones. Pecuniary harms include damage to one's property or earnings capacity or the creation of legal liabilities; nonpecuniary harms include bodily harm, emotional distress, humiliation, fear and anxiety, loss of companionship, loss of freedom, distress caused by mistreatment of a third person or a corpse, and so on.¹⁸

Now, compensation of the sort lawyers have in view will come in a pecuniary form, as monetary damage awards or other cash payments. Hence, pecuniary compensation for pecuniary losses would constitute what I have called compensation: the replacement of like with like. Compensation of a pecuniary sort for losses which themselves were nonpecuniary seems to constitute compensation: the substitution of one sort of pleasure for another.

One good indicator of the balance of compensation, to compensation, in our existing compensation policies, then, is the extent to which they attempt to compensate for pecuniary versus nonpecuniary losses.¹⁹ In practice, the former typically involves payments to replace

inction between "equivalence" compensation and "substitute (or solace)" compensation, although the examples offered there tend to blur the distinction.

¹⁸ Prosser and Wade 1979, secs. 905-6.

¹⁹ Another indicator is the way in which courts order "specific performance" of contractual duties where "the existence of sentimental associations and esthetic interests" or "the difficulty, inconvenience or impossibility of obtaining a dupli-

lost earnings or to cover extra expenses associated with injuries or disabilities, whereas the latter typically involves payments compensating for "pain and suffering" or the "loss of faculties or amenities."²⁰

Perhaps the most comprehensive survey of compensation policies is the one carried out by the Oxford Socio-Legal Studies Centre in the late 1970s. Eighteen categories of financial support to U.K. victims of illness or injury were studied.²¹ Of these, only four (or perhaps five) offer any provision at all for pain-and-suffering or loss-of-faculties payments.²² Summarizing these findings, the Oxford team writes,

Most benefits . . . give priority to meeting either the loss of income or the reimbursement of the extra expenses incurred by disabled people. A few - damages, criminal injuries compensation, the disablement benefit for industrial injuries, and war pensions - do provide some money to assuage suffering or to give an alternative pleasure where the . . . victim can no longer enjoy a particular activity. But this type of loss is covered by social security only in exceptional cases, and few people take advantage of the opportunities to buy private insurance to cover against it.²³

Furthermore, among those programs offering pecuniary compensation for nonpecuniary losses, only one (tort law) provides substantial sums to large numbers of people in many jurisdictions. "Personal accident insurance policies (rare enough in themselves) are usually limited to medical expenses or income losses; and though small disability payments are often made under comprehensive road traffic insurance policies, they rarely exceed £500 for severe disablement, with lesser sums for other cases."²⁴ Compensation for pain and suffering or loss of faculties associated with war injuries, industrial injuries or criminal injuries is obviously available to only very small

cate or substantial equivalent of the promised performance" means "the remedy in [monetary] damages would not be adequate" (Williston 1932, secs. 358[1] and 361[b-c]).

²⁰ Atiyah 1980, pp. 535-9; Atiyah and Cane 1987, pp. 473-6. D. Harris 1974, p. 48. Prosser and Wade 1979, sec. 903, comment a.

²¹ These include two types of damages (damages at common law, as modified by statute; criminal injuries compensation), ten types of social security income support (industrial injury benefit; disablement benefit, and special hardship allowances and unemployment supplements thereto; war pensions and associated special allowances; sickness benefit; invalidity benefit; noncontributory invalidity pension; invalid care allowance; supplementary benefit); four types of social security expense payment (attendance allowance; constant attendance allowance; mobility allowance; the family fund) and two types of private provision (sick pay from employers; private insurance). See Harris et al. 1984, pp. 4-12.

²² These are criminal injuries compensation, disablement benefit, war pensions and (often) private personal accident insurance (Harris et al. 1984, pp. 4-12).

²³ Harris et al. 1984, p. 15.

²⁴ Atiyah 1980, p. 537; Atiyah and Cane 1987, p. 475.

numbers of people injured in very particular circumstances. Even then, the pain-and-suffering or loss-of-faculty component in the award (certainly as compared to the loss-of-earnings component) is typically quite small, amounting to no more than \$500 to \$1,000.²⁵ Tort law, although notionally more generous, in practice often offers little more.²⁶

Sums like these can hardly pretend fully to "make up" for serious bodily harm. They are instead token payments. As with "nominal damages" in tort law, the sums involved are not "utterly derisory"; but pretty clearly, the principal value of the awards is meant to be symbolic. The aim, in Atiyah's terms, is surely to provide "solace" rather than "substitutes."²⁷

Thus it would seem that monetary payments principally serve to replace monetary losses. The vast majority of compensation programs doling out pecuniary awards do not even try to compensate for non-pecuniary losses at all. Those few that do tend, in practice, to make only token gestures along such lines. That strong preference for replacing like with like, money with money, would seem to betray a preference for compensation, over compensation.

The same pattern reappears when we look more deeply at the way in which compensation schemes characteristically function. We provide the blind with talking books, readers and audible street-crossing signals. We provide the wheelchair-bound with access ramps to public buildings. We provide invalids with home help (or an Invalid Care Allowance, to allow them to hire it), and the lame with transport (or a Mobility Allowance, to allow them to acquire it), and the disabled with rehabilitation and retraining.²⁸

All those things are by way of compensation, – improving people's lives in broadly the same respects as some accident, injury or disability has worsened them. What we typically do *not* do is to offer compensation, compensating people in one realm for losses suffered in some other realm entirely. Monogamous societies do not, typically, make an exception to allow a blind man to take two wives. That might make

²⁵ Ison 1980, p. 65; Nonet 1969, pp. 20–5; Enker 1974, vol. 2, p. 131; Elias 1983, pp. 33, 151–7.

²⁶ In one study of out-of-court settlements, the "mean sum for non-pecuniary losses such as pain and suffering was £ 973," which does indeed seem low given that nearly 40 percent of cases involved permanent disability of some form or another (Harris et al. 1984, p. 90).

²⁷ Williams and Hepple 1984, pp. 57–8. Atiyah 1980, p. 537; Atiyah and Cane 1987, p. 475. See similarly Enker 1974, vol. 2, p. 131.

²⁸ Haveeman et al. 1984, pp. 45–6; both what they term "ameliorative" and "corrective" responses would fall within my larger category of compensation. See details of the programs in Harris et al. 1984, chap. 1; and Atiyah 1980, chap. 16; Atiyah and Cane 1987, chap. 16.

him better off in some global sense. But it would be deemed inappropriate, having nothing to do with his blindness.

III. THE STRUCTURE OF PREFERENCES AND THE POSSIBILITY OF COMPENSATION

Modern welfare economists no doubt would, on the face of things, find this preoccupation with compensation, baffling. From their perspective, the point of compensation is merely to leave people as well off as we found them. If indifference curves are conceptualized as connecting points representing different bundles of goods that a person regards as as good as each other, then being compensated must surely just be a matter of ending up on the *same indifference curve* afterward as before. There is no need to restore someone to the *same point* on that indifference curve (i.e., to restore exactly the same bundle of goods to him) since, *ex hypothesi*, he is indifferent between all alternative points on the same curve.⁹ For the welfare economist, the choice between compensation, and compensation, all comes down to cost; and if in practice it proves cheaper to make losses up to people in some way other than by restoring things like those they have lost (as typically it will),¹⁰ then compensation, is from the welfare economist's perspective decisively to be preferred.

What underlies welfare economists' insensitivity to the distinction between compensation, and compensation, is their studied indifference to the deeper structure of people's preferences. Within conventional consumer theory, everything is presumed to substitute for everything else at the margins.¹¹

Now, even within economics there is a growing band challenging this presumption. Georgescu-Roegen wryly observes that "bread cannot save someone from dying of thirst, . . . living in a luxurious palace does not constitute a substitute for food, etc."¹² Or as Lancaster says, there must be something about margarine that makes it a good sub-

⁹ Economists tend to assume that any "nonpecuniary" losses can be compensated just by adding more money to the sum of compensation paid (Kaldor 1939, p. 551 n. 1; Culyer 1974, p. 22).

¹⁰ Obviously, getting him back to the same indifference curve (compensation,) can never cost more than getting him back to some particular point on that curve (compensation,); often it will cost less.

¹¹ Pareto [1927]1971, pp. 182-6; Wicksteed 1933, pp. 152-3, 360-1. Cf. Sen (1977a) and Nozick's (1968, pp. 33ff.) discussion of the various different ways in which values might "override, outweigh, neutralize, weaken, dissolve, destroy, invalidate, preclude or nullify" one another, which suggests a similarly complicated structure.

¹² Georgescu-Roegen 1954, p. 516.

stitute for butter but a bad substitute for a Chevrolet. Building on such observations, Lancaster goes on to offer his New Consumer Theory,

breaking away from the traditional approach that goods are the direct objects of utility, and instead supposing that it is the [objective] properties or characteristics of the goods from which utility is derived. . . . Utility or preference orderings . . . rank collections of characteristics and only rank collections of goods indirectly through the characteristics that they possess.³¹

In Sen's terms, "commodities" are valued not only in their own right but rather by virtue of the "capabilities" that they bestow.³² In short, goods have certain objectively defined capacities to serve our subjectively defined ends.

The particular importance of this model for the present argument lies in its analysis of the way in which goods can substitute for one another. One thing is a good substitute for another if, however different it might otherwise be, it has the same objective capacity to promote exactly the same end as does the other. In Lancaster's terms, two goods are "perfect substitutes" if they present exactly the same "characteristics" in exactly the same proportions; they constitute "close substitutes" if the associated characteristics-bundles are substantially similar. In Sen's terms, they are good substitutes insofar as they promote the same capacities.³³ Thus, objects that are otherwise very different – as are trains and cars (ask any engineer) or butter and margarine (ask any chemist) – might nonetheless constitute close substitutes for one another, insofar as they present the same deeper Lancasterian "characteristic" or promote the same "capabilities" or, in layman's terms that connote almost the same thing, serve the same ends.

For many things, there *are* close substitutes. Production-line manufacture being what it is, one Ford Fiesta is to all intents and purposes just like another. So, unless you happen to form sentimental attachments to your automobiles, you can be fully compensated in the first sense as well as the second for the loss of one Ford Fiesta by being given another. One dollar bill is much like another. So, unless you attach particular importance to how you came by it (e.g., it was the first dollar you ever earned, or it was given to you by your grandmother before she died), you can be fully compensated in the first sense as well as the second for the loss of one dollar by being given another. And so on.

There are many things, however, for which there are *no* close sub-

³¹ Lancaster 1966, p. 133; 1971.

³² Lancaster 1966, p. 144.

³³ Lancaster 1966, p. 144. Sen 1985a.

stitutes. One rich source of examples concerns personal integrity. Both bodily integrity and moral integrity are the sorts of things that, once lost, are largely irreplaceable. Other examples concern goods which are valued on account of their histories. Works of art, keepsakes, historical landmarks and natural wonders are all irreplaceable insofar as what we value about them is intrinsically bound up with the history of their creation. That is what makes facsimiles, which are otherwise identical to their originals, mere "fakes."⁶ There being no close substitutes for objects that are irreplaceable, it is impossible to compensate people in the first sense should these things be lost. All we can do is to compensate them in the second sense, offering them goods with different characteristics, speaking to altogether different ones of their desires, and yielding altogether different satisfactions.

The welfare economist's case for ignoring any distinction between the two kinds of compensation, sketched in the opening paragraph of this section, was that "indifference is indifference; it does not matter where compensation puts you on an indifference curve, just so long as you are restored to the *same* curve." Recasting the argument of this section into those terms, we have seen that indifference is not all of a cloth. There are, in fact, two kinds of indifference, corresponding to the two kinds of compensation.

In the form of indifference that parallels compensation, we might be indifferent, between two options because they are equivalent ways of achieving the same goal. We might be indifferent, between the high road and the low road because they both get us to the same destination in the same time and with the same effort. In the form that parallels compensation, we might be indifferent, between options because they are ways of achieving equivalently good goals. We might, for example, be indifferent, between the Glasgow road and the Edinburgh road because both cities offer amusements which, however different, are equally amusing. Economists, in their continuing quest to "extract the minimum of results from the minimum of assumptions,"⁷ use the same curve to represent both fundamentally different phenomena.

IV. THE SUPERIORITY OF COMPENSATION,

With this apparatus in hand, we can now return to the question of how it can be wrong for the state to do certain things to people, even if it compensates them for their losses. The short answer, foreshad-

⁶ On "irreplaceable assets," see Goodin 1983; 1982a; 1982b, pp. 120-1, 157-8, 181-3. On "fakes," see Goodin 1992a, chap. 2, building on Sagoff 1978 and Elliot 1982.

⁷ Lancaster 1966, p. 132.

owed at the outset, is that the compensation is inadequate to legitimate the policy, because it is of the wrong kind. The cases where compensation is inadequate to legitimate policies, I submit, are cases where there is something irreplaceable at stake. Since there are no close substitutes for the things people would lose, the state could compensate them only in the weaker, second sense. And that is just not good enough.

Why is that not good enough? After all, something might be irreplaceable without being of infinite value. Each oil painting is, in some sense or another, an utterly irreplaceable "one of a kind." That, however, does not stop artists (even rich ones, who are not in any sense acting under duress) from selling their works. The same seems to be true for a wide variety of other things that we would regard as irreplaceable. There is usually *some* price such that people would be induced to part with them.

But it is one thing for someone, in exchange for something else altogether, voluntarily to part with some thing that is irreplaceable.³⁸ It is quite another for the state compulsorily to force that trade.

The way that compensation works to legitimate public policies is by removing any distributional objections to the consequences of those policies. That is clearly the role economists see it playing. If gainers actually compensate losers and still have some gains left over, then the policy constitutes a Paretian improvement in that someone wins and no one loses. If gainers hypothetically could compensate losers and still have some gains left over, then at least that shows we could have neutralized the distributional effects of the policy and still shown a profit; that we refuse to do so is itself a distributional decision.³⁹

There is nothing peculiarly economic in viewing compensation in this way. Lawyers and courts of law have long taken a similar view of it.⁴⁰ What is peculiarly economic is the way of putting the point. In explaining how compensation removes distributional objections, the economist would typically say something along these lines: "If everyone is as well off as he was before the policy was instituted, then no one has any grounds for complaint."

That way of putting the point, however, focuses on *interpersonal*

³⁸ That slides over the question of what constitutes a "coercive offer." If someone has no choice but to accept the offer - because otherwise he will die, for example - then the offer counts as "coercive," whatever its external form. The same may be true if the sum offered is very substantial compared to the person's present holdings (e.g., imagine an underpaid clerk being offered \$1 million in exchange for his little finger).

³⁹ That is its role in discussions both theoretical (Kaldor 1939; Hicks 1939) and applied (Cordes 1979; Cordes and Goldfarb 1983; Hochman 1974; Tullock 1978).

⁴⁰ Michelman 1967, p. 1168. Ackerman 1977. Tribe 1978, chap. 9.

redistributions while ignoring *intrapersonal* ones. As shown in Section III, people's preferences and goals are not one undifferentiated mass. Rather, they fall into several distinct, subjectively defined categories. To guarantee the distributive-neutrality of our policies under those circumstances, it is not enough that people be left globally *as well off* as we found them. We must furthermore make sure they are left exactly *as we found them*. The former consideration speaks to interpersonal distributions, the latter to intrapersonal ones. It would be wrong, to the same extent and for the same reasons, for the state peremptorily to redistribute priorities between the goals and projects that constitute one person's own life as it would be to redistribute resources between the goals and projects that constitute different people's lives.⁴¹

Compensation, where it is possible, successfully avoids both sorts of distributional objection. Where they are given close substitutes (as defined previously) for what they have lost, people are not only as well off as before but also in exactly the same position with respect to exactly the same goals as before. All that has changed is the means by which those goals are to be pursued.⁴² Where no close substitutes are available for what has been lost – where compensation, alone is possible – some amount of intrapersonal redistribution is inevitable. People might be as well off as before, but they will be differently off. They will have been forced to shift their goals, and not just their means of achieving their goals. Thus, compensation, erases all distributional objections to policies, whereas compensation, erases only half of them. That explains the superiority of the first sort of compensation over the second. That explains why compensation, is just not good enough to legitimize certain sorts of policies.

(Again, I should emphasize that distributive neutrality is neither a necessary nor a sufficient criterion of a legitimate policy, from a broader perspective. As I said at the outset, all kinds of state action are perfectly permissible without any compensation whatsoever. My point here is a much narrower one. The only way compensation can do anything at all to render legitimate otherwise illegitimate policies is by removing distributional objections to them; and compensation, by itself, can do only half that job.)

⁴¹ Economists might protest that the two are disanalogous. In the interpersonal case, the distributional objection is that someone has been harmed; in the intrapersonal case, no one has. But that is true only if "having been harmed" is completely analyzable in terms of "having been shifted to a lower indifference curve," which I am here denying.

⁴² That in itself might constitute an interference with the course people have chosen for their lives, but it obviously constitutes much less of one than forcibly shifting them onto different goals altogether.

There are two independent ways of explaining what, exactly, is wrong with imposing on people such intrapersonal redistributions, forcibly shifting them from one set of plans and projects to another. The first has to do with the value of "coherence and unity" in a person's life.⁴³ Critics of classical utilitarianism have made much of the objection that it requires us to lead an incoherent life: fifteen minutes collecting for Oxfam, three hours as a nurse, twenty minutes as an investment banker, two hours shearing sheep, etc. Such a life, maximize social utility though it may, in some deeper sense adds up to nothing in the end.⁴⁴ One way of capturing this thought is to say that you can be either a saint or a sinner, but that there is no point in being a saint and sinner on alternate days.

As it stands, this is a perfectionist objection to forced intrapersonal redistributions between a person's plans and projects. That is to say, a life characterized by more coherence at any moment in time and by more stability across time is a "better" life, by some *external* criterion.⁴⁵ Of course, it is also a more satisfying life by the internal criteria that most people use in deciding what makes their own lives satisfying. But some people might happen to prefer a less coherent life to a more coherent one – regarding "coherence" as a straitjacket constraining creativity, or whatever. Given this potential divergence, perfectionist arguments based on the objective goodness of a more coherent and unified life are potentially open to powerful antipaternalist rejoinders.

Remember, though, that the objection here in view is to *forced* shifts between a person's plans and projects. If someone freely chooses to adopt and abandon projects willy-nilly, that would be one thing. Even if we suppose that would be a less good life, by some external standard, we might nonetheless suppose that he should be allowed to lead his own life as he pleases. But for someone to be forced, by some external agency, to drop one project and take up another (even one that he would himself regard as an equally good project) is something else altogether. Far from endorsing that policy, the antipaternalist argument firmly condemns it.

Second is the logically quite separate argument, from "autonomy," against forced intrapersonal redistributions between a person's plans and projects. It is, after all, a central tenet of the liberal ethos that "respecting" people means taking them as we find them. It is important, in those terms, that people should be free to choose their own

⁴³ Compensation has long been justified as providing stability, and hence coherence, in people's lives (Sidgwick 1891, pp. 179–80; Tullock 1978, p. 54; MacCormick 1982, p. 214). This theme is further developed in Chapter 12.

⁴⁴ B. Williams 1973a, pp. 108–18; 1981, chap. 1.

⁴⁵ Nozick 1981, pp. 403–51. Wollheim 1984.

life plans for themselves; and it is equally important, in those terms, that once those choices have been made other people should respect them.⁴⁶

Modern welfare economics grasps this point, albeit imperfectly. There, "the criterion used for [specifying] an increase in an individual's welfare [is] that he is in a chosen position."⁴⁷ But surely taking people as we found them means respecting people's *actual* choices – ones that they *really made*, rather than ones they might have made in some counterfactual world that never has (and perhaps never will have) existed.⁴⁸ What we are supposed to be respecting is people's choices, not their disembodied preference orderings. It would be flatly contrary to the fundamental ethos of liberal welfare economics to force people to consummate Pareto-optimal deals, or to make such trades on their behalf without their permission.⁴⁹ Suppose someone has contrived to sell my house out from under me, without my consent. Surely it would not suffice for him to reply to my protests that he got an exceptionally good price for it and that, despite the fact it was not for sale, I certainly would have agreed to sell it for that price if only he had been able to contact me. Whether or not I would have agreed, the point remains that I did not. By virtue of that fact alone, my autonomy has been violated.⁵⁰

Means-replacing compensation, respects both of these values, whereas ends-displacing compensation, respects neither. Providing people with alternative means to the same ends (compensation,) allows them to pursue the same, self-selected goals as before. That they are the "same" ensures unity and coherence; that they are "self-selected" ensures autonomy.⁵¹ Compensation, in contrast, might leave people "as well off as before," in some sense or another, but it forces them to pursue different goals than before. That they are different compromises unity and coherence; that they are forced compromises autonomy.

Compensation, in effect, forcibly pushes people along their indif-

⁴⁶ O'Neill 1985. Respecting choices for the sake of the dignity of making them (Goodin 1982b, chap. 5) similarly suggests that compensation, is an inadequate substitute for compensation. Mucking around with a person's life plans, forcibly shifting him from one goal set to another (even if it is, from his own point of view, an equally good set of goals) is hardly the way to preserve the person's self-respect or self-image.

⁴⁷ Little 1957, p. 37.

⁴⁸ R. Dworkin 1977, chap. 6; [1980]1985, pp. 275–80.

⁴⁹ Barry 1986, pp. 11, 41. Sen 1977a, p. 93.

⁵⁰ Broome 1978a, p. 316. Kleinig 1982. Calabresi and Melamed 1972, p. 1126.

⁵¹ Compensation, may be more intrusive – rehabilitation involves more of an intrusion into a person's life than cash compensation, for example. But that only goes to show that autonomy is not simply a matter of nonintrusiveness.

ference curves. The fact that a person remains on the same indifference curve means that, *ab initio*, he would have been equally prepared to accept either option, either his previous bundle of goods or his new bundle.³⁹ He *would* have been – but, as a matter of personal history, he did not (his life has gone down a different track, now); and, as a matter of public morality, no one ever asked (he did not consent to the change). Morally, both those facts are vital. For those reasons, when a new bundle of goods is simply foisted upon people in compensation, whether or not it is an equally good bundle is simply irrelevant.

V. POLICY IMPLICATIONS

Insofar as we are counting on compensation to right what would otherwise constitute wrongful inflicting of harms upon people, we must respect the following precepts that follow from the arguments developed thus far.

(1) *Prevention is better than compensation, where it is an irreplaceable object that would be lost.*

The logic of this proposition is simple. If something irreplaceable is lost, only the weaker form of compensation, would be possible. People would be as well off but differently off than before. If the loss is prevented, however, that would leave them in exactly the same position as before, still in possession of the irreplaceable object itself.⁴⁰

This explains the differential, noted by several economists, between how much people are prepared to spend to protect certain things and how much they are prepared to insure them for. Zeckhauser offers the compelling example of a woman facing the risk of breast cancer. Imagine she is willing to spend \$5,000 for medical treatment to reduce the risk of cancer from 10 percent to 5 percent. That implies that the value of a healthy breast to her is \$100,000. Suppose now she is offered insurance at the rate of \$20 of coverage per dollar's premium. Does it necessarily follow that she will pay \$5,000 more to cover the full \$100,000 that the breast is worth to her? Zeckhauser concludes that it

³⁹ In Thomson's (1986, chap. 10) terms, he would have been willing, *ex ante* of his rights being violated, to sell the violator permission in exchange for the price now being paid in *ex post* compensation for their violation.

⁴⁰ Atiyah 1980, p. 8; Atiyah and Cane 1987, p. 7. Irreplaceability apart, psychometric evidence shows that people attach more value to avoiding loss of what they already have than they do to securing symmetrically large gains (Kahneman and Tversky 1979; Knetsch and Sinden 1984; Gregory and McDaniels 1987); on the implications of this fact for the economic analysis of the law, see Mishan 1967: 1971, pp. 19ff.

does not. Since the insurance money would not restore the breast, "it would be quite rational for her to insure no more than the medical expenses" of the mastectomy. Similarly, when a Constable painting valued at £100,000 turned out, after having been stolen, to have been insured for only £2,000, the vicar explained: "We never had any intention of selling it and we could never replace it so there wasn't any point in insuring it for its full value."⁵⁴

This principle is also reflected in certain practices of the courts. Ordinarily, the courts let people do as they will; and they order tort damages *after* the fact if people have, in the end, caused others some harm. Sometimes, however, the "nature of the interests" that stand to be harmed is such that damages would be a "relatively inadequate remedy." Where the interests that would be "harmed by tortious conduct are so remote from the marketplace that . . . it is idle to speak of their compensation in terms of money," courts will not wait until after a tort has been committed. Instead, they will issue an injunction designed to prevent the tort from ever occurring.⁵⁵

Finally, notice that much that presents itself as compensation policy might just be an oblique form of prevention policy. This is so because, in many realms of compensation policy, the compensation would have to be paid (in whole or in part, directly or indirectly) by the persons responsible for causing the damage. Tort law is the clearest example of this; a weaker one might be workmen's compensation, where the employer's contributions are uninsurable or where the premium paid for such insurance varies according to the number of claims the insured has lodged. Since the risk of incurring such expenses would presumably serve to deter people from actions likely to harm others, compensation policies in this way might double as prevention policies. It is hard to discern what the balance might be as between these two very different aims in present compensation schemes. But the prevention rationale clearly does explain what otherwise appear as anomalies in present policies, such as the awarding of compensation for "loss of amenities" to a person who, through severe brain damage, has been rendered "totally insensitive to his loss." Here the deterrent/prevention rationale is clearly controlling. The principle at work is simply that "it should not be cheaper to kill than to maim, and, further, it should not be cheaper to injure a person so severely that he is incapable of obtaining any enjoyment from a sum awarded to him as compensation than to injure him less severely."⁵⁶ That argument has

⁵⁴ Zeckhauser 1975, p. 454. S. Reeve, *Colchester Evening Gazette*, October 23, 1985, p. 3. See more generally Cook and Graham 1977.

⁵⁵ Prosser and Wade 1979, secs. 936, 944 and 944 comment b. G. Williams and Hepple 1984, pp. 68-73.

⁵⁶ G. Williams and Hepple 1984, p. 83. Atiyah and Cane 1987, pp. 187-8. On de-

nothing to do with the adequacy of compensation for the victim, and everything to do with the adequacy of the deterrent for the tortfeasor.

(2) *Where a lost object is replaceable, the compensation offered should include the closest possible substitute for that which has been lost.*

The aim of that form of compensation which can legitimize otherwise illegitimate state action – compensation, – is to allow people to remain in exactly the same position with respect to exactly the same ends as before the damage occurred. The goal is to make sure that means can be replaced without ends being displaced. The more nearly perfect the replacement (the better the substitute) that is being offered, the more nearly this goal of compensation, has been accomplished.

This principle goes some way toward explaining why we are relatively comfortable in compensating people for losses that can be truly said to have some "fair market value." The advantage usually claimed for this class of cases is that here we can unambiguously fix a fair (i.e., market) price on the losses.⁵⁷ But that, I think, is the smaller part of the story. The real advantage in such cases lies, I think, not in the fact that there is a *market price* for those things which are marketed. It lies instead in the fact that there is a *market* in those things which are marketed. That is to say, people can take the money they receive in compensation, go out into the marketplace, and buy another object just like the one they have lost.

This principle also explains the emphasis upon *rehabilitation* in so many of our actual compensation policies, detailed in Section II. Rehabilitation, understood literally, consists in restoring lost functioning of that which has been damaged; understood metaphorically, it consists in substituting for that which has been damaged something that will perform much the same function. Occupational therapy is an example of the former, prosthetic devices of the latter.⁵⁸

This emphasis upon rehabilitation also goes some way toward explaining why public policy should so often strive to aid the injured (and disabled, in particular) as a group rather than as individuals. As Donald Harris observes,

Handicapped people are usually dependent on governmental or community projects to provide them with specially-adapted housing or transport, parking and recreational facilities, access to buildings open to the public such as museums, theatres, cinemas, etc. The common law

terrence and accident prevention more generally, see Atiyah and Cane 1987, chap. 24 and Calabresi 1970.

⁵⁷ Corwin 1978, p. 402. Calabresi and Melamed 1972, p. 1108. Prosser and Wade 1979, sec. 903 comment a.

⁵⁸ Haveman et al. 1984, chap. 4.

notion of giving the individual his own sum of money to find his own facilities on an individual basis is not realistic in the modern world.⁶⁰

Giving someone who has been crippled monetary damages does not help him up the stairs to the City Council chambers, whose meetings he used regularly to attend. Building him a wheelchair ramp does. "The importance of these facts is that they suggest that public expenditure of money to overcome difficulties of this kind may be a higher priority than more private compensation for disabilities as such."⁶¹

(3) *People should be compensated as best they can for irreplaceable objects once lost, but that does nothing to legitimize policies deliberately inflicting those losses.*

Sometimes people suffer irreparable losses, despite our best efforts at preventing them. Or sometimes we find ourselves inflicting irreparable losses as part and parcel of some policy that is independently legitimized whether or not compensation is paid. Once irreplaceable objects have been lost, compensation, is the only possible remedy. It is a very inadequate remedy, to be sure: *ex hypothesi*, there are no close substitutes available. Still, inadequate though it may be, compensation, is surely better than nothing. There can be little doubt that it should be paid.⁶²

We must, however, be very clear as to what its payment might accomplish. Payment of compensation in the strong sense – compensation, – can right wrongs fully and completely legitimize our loss-inflicting course of conduct. Payment of compensation in the weak sense – compensation, – cannot. Insofar as losses are irreparable, compensation is necessarily inadequate. And insofar as compensation, is thus inadequate, so too is the plea that "compensation has been (will be, could be) provided" inadequate to excuse a loss-inflicting course of action that would otherwise be illegitimate.

The distinction I have in mind here can best be evoked by examples from criminal injuries compensation policies. It is one thing to pay the widow of a soldier killed by IRA snipers £100,000 in compensation after the fact; it is quite another to use that sum in deciding ahead of time whether or not to buy soldiers flak jackets that would save their lives.⁶³ Or, for another example, it is one thing to decide that we

⁶⁰ D. Harris 1974, p. 48.

⁶¹ Atiyah and Cane 1987, pp. 379ff.

⁶² Calabresi and Melamed 1972.

⁶³ Cf. Tullock 1978, pp. 53-4. This leaves open the question of whether it is permissible to impose or incur mere risks of such losses; for diverse views on this, see Nozick 1974, pp. 82ff.; Goodin 1982b, pp. 157-8; and Thomson 1986, chap. 11.

should pay rape victims £1,000 in compensation; but it would be quite another to decide that it would be cheaper to pay off the two victims that will predictably get raped in some particularly dark street than to install a £3,000 lighting system. That compensation of this sort is inadequate does not mean that it should not be paid at all. But it does mean that it should not be counted on to right all the wrong. Prevention is still the best policy.

VI. CONCLUSION

For our conclusion, let us return to that classic cautionary tale concerning economism and public policy, the Roskill Commission. Among the things that needed to be calculated in reckoning the costs and benefits of a third London airport were the losses to homeowners who would be displaced. Just reckoning the value of a house at its market price obviously understates the true value of the house to the householder. After all, he declines to sell his house on the free market at the market price. What right do we then have to assume that he would be fully compensated for its loss by the same price he has already rejected when it is compulsorily purchased by the government? So the Roskill Commission set about surveying residents, asking, "What price would be just high enough to compensate you for leaving this house (flat) and moving to another area?"

The striking thing about this survey was that 8 percent of respondents said that they would not move at any price. Now, as Mishan says, "it may be that a good interviewer would have elicited a finite sum . . . - perhaps £50,000? or £5 million?" But, as he goes on to say, "it is not altogether inconceivable that for some older, or unworldly, people all that [money] could buy for them would not suffice as compensation for having to live elsewhere."⁶

Presumably few people would be so silly as to deny that with £5 million in compensation they would, in some sense, be better off moving out of their £5,000 house and living elsewhere. What these respondents would surely have said is not that they are better off, but rather that no amount of money can replace lost friends and the like. In my terms, it is the impossibility of compensation, not the inadequacy of compensation, that was at issue here.

This, I dare say, is a common pattern. Most policies will probably run up against at least 8 percent of losers who feel hard done by in some such way. That is not to say that we should not carry forth with the policy. There are all sorts of reasons for and against building a third London airport; the uncompensatable, loss of displaced residents

⁶ Mishan 1970, pp. 462-3.

is just one among many, and on balance we may well decide that it is best to go ahead with the policy. What we cannot say, however, is that since losers will (or could) be compensated, they have no grounds for complaint.⁶⁴

⁶⁴ Earlier versions were presented at the Universities of Arizona, Chicago, Georgetown, Göteborg, Maryland, Pennsylvania, Stockholm, Uppsala and York and the Australian National University. I am grateful for the comments, then and later, of Geoff Brennan, John Broome, John Dryzek, Jim Griffin, Russell Hardin, Sheldon Leader, Julian Le Grand, Keith Lehrer, Howard Margolis, Phillip Pettit and Gordon Tullock.

Ensuring social security

Stabilizing expectations

The political history of welfare state institutions is one thing, their ethical justification another. The collective intentions of the many and varied actors who had a hand in shaping such institutions were seldom simple and rarely altogether noble. For present purposes, let us leave actual intentions to one side, however. Let us ask, instead, whether there are any good moral reasons (whether or not they were actually the founders' reasons) for the institutions that they have bequeathed to us. Putting the point another way, were we constructing our social institutions *de novo*, would we have good moral grounds for including a welfare state with these particular characteristics among them? It is in this spirit that I here inquire into the practice of paying certain social benefits in an earnings-related form.

Conventional moral wisdom has long held that the welfare state is justified principally as a device to benefit the poor. There may, of course, be perfectly respectable reasons for its benefiting not only the poor.¹ Pragmatists, reflecting upon the realities of political power and economic behavior, may counsel that the price to be paid for programs that benefit the poor is to allow the nonpoor to cream off some of those benefits, too. Idealists, reflecting upon the value of community solidarity, may offer some more high-minded reasons for wishing the welfare state to benefit rich and poor alike. Whoever else is meant to benefit, and why ever they are meant to benefit, though, it is standardly presumed that the welfare state's first concern must always be with the poor. They, always and ever, are the principal direct intended beneficiaries of welfare state programs. Or so goes the standard story told in moral justification of such schemes.

Given that understanding of the welfare state, we would be hard pressed to explain why welfare programs should sometimes actually

¹ Goodin and Le Grand 1987.

- and, indeed, intentionally - pay more to the better off. Yet that is precisely what is done through earnings-related benefits built into many social welfare programs. The more you earn, the more you will receive when you cease to earn, in the form of old-age pensions, unemployment benefits, sickness and disability benefits, maternity benefits, and so on down the list.

There may, of course, be no good ethical justification for that practice. Perhaps it is nothing more than a politically necessary (or, still worse, unnecessary) sop to the better off to buy their support for pro-poor programs. If we can find no good moral justification for earnings-related benefits, that will be the conclusion with which we are left. But we ought not leap to that conclusion precipitously, ahead of seeing whether some ethical justification can indeed be found.

Here I examine the administrative details of such programs in an attempt to tease out various possible grounds for linking benefits to earnings in these ways. I then go on to explore to what extent the rationales thus uncovered provide leverage on larger questions of how best to justify the welfare state as a whole. The model I end up defending traces one of the welfare state's principal justifications to its role in providing "stability" or "security" for citizens. By smoothing out the peaks and troughs in their earnings patterns, it provides limited protection for their market-based expectations about future income streams. Morally speaking, that is a modest but nonetheless worthy goal.

I

In social security law, payment of benefits is conditional in various ways on various factors. Some of the conditions serve to restrict who is entitled to receive benefits. Other of the conditions serve to determine how much benefit they are entitled to receive.

The former sorts of restrictions are the more familiar. Broadly speaking, these conditions fall into two types. First, there are programs of "categorical assistance" (old-age pensions, disability benefits, workmen's compensation, unemployment assistance, widows' benefits, and such like) providing benefits only to those who are deprived or disadvantaged in some specific respect or for some specific reason. Second, there are programs of "general assistance" providing benefits to anyone who is poor, regardless of reason. Eligibility for general assistance is here conditional simply upon a test of the claimant's income or assets or both.

The relative merits of alternative ways of determining who should be eligible for social welfare benefits are much discussed, among mor-

alists and policy-makers alike.¹ The further question, no less important but much less widely discussed, concerns how we should set about determining how much benefit ought to be paid to each eligible claimant.

In programs of general assistance, the two questions tend to be answered in the same way. The same sort of means test that is used to determine eligibility in the first instance is used to determine benefit levels. Claimants are eligible for this form of assistance because their income or assets fall below a certain threshold, and they are eligible for more assistance the further they fall below that threshold.

Under such rules, the aim is to ensure that only the poor get assistance and that they get more assistance the poorer they are. Those aspirations may not always be completely realized, of course,² but at least the motivation underlying them is readily comprehensible. No matter whether one's aims were to relieve poverty or to reduce social inequality, a good way to further such aims would be to target social benefits on those people with less than median income or assets and to pay them more the further their incomes or assets fell below that median.

In programs of categorical assistance, in contrast, eligibility and benefit levels are decided according to different criteria. Eligibility for benefit, in the first place, is predicated upon a claimant's fitting the description set down in the categories of persons entitled to that benefit. Once that categorical hurdle is crossed, however, the level of benefit to which claimants are entitled is predicated – either in whole or in very larger part – upon the claimant's past earnings. Details of social security law naturally vary from country to country. But it is striking how nearly universal this basic phenomenon actually is. Judging from the U.S. Social Security Administration's 129-nation synopsis of social security throughout the world, in virtually every major categorical assistance program in virtually every nation of the world benefits are strongly and positively linked to previous earnings.⁴

¹ The standard reason for focusing on those particular categories of distress, and for treating people suffering those forms better than poor people in general, involves notions of the "deserving poor." Recipients of categorical assistance are presumed to be poor through no fault of their own, whereas those eligible only for noncategorical general assistance are presumed to be somehow to blame for their plight. Whether that is empirically true or ought morally matter is questionable (Goodin 1988, chap. 10). But as an analysis of the bases of social policy, that seems indisputable.

² Goodin and Le Grand 1987; Le Grand 1982.

⁴ U.S. DHEW 1978, cols. 5 and 6 in each of the country summaries. "Virtually all programs" because, while benefits are almost invariably earnings related in four classes of programs (old age, invalidity and survivor programs; unemployment benefit; sickness and maternity programs; work injury programs), they never are in a fifth (family or child allowances). "Virtually every nation" because a few countries pay flat-rate benefits under some programs, though even they typically pay earnings-related supplements under at least some of their other programs of

Notice the curious consequences of this practice. Where benefit levels are earnings-related in this way, the more that you earned during some "qualifying period" prior to the commencement of receipt of benefits the larger the absolute benefit that you will receive in consequence. Under such a rule, less – not more – is paid to those who are and have always been poor. Furthermore, the poorer they have historically been, the less benefit they will receive directly in consequence of that fact.

The pragmatic motivation underlying such practices is, perhaps, clear enough. Insofar as it is economically or politically necessary to bribe the nonpoor into supporting social welfare programs by letting them benefit too, the bribe naturally needs to be larger, the richer the person who is to be bribed. Much in the historical record suggests that just such cynical logic led to the development of earnings-related components of welfare systems.¹

If we are looking for principled rather than baldly pragmatic rationales, however, these consequences of earnings-related benefit structures are far harder to justify. Even moderately large programs of earnings-related benefits can do little to alleviate poverty, and even less to narrow social inequalities. They necessarily frustrate communitarian attempts at producing social solidarity and fraternity by treating everyone – rich and poor – alike. In short, linking benefits to prior earnings has obvious and inevitable effects that run clearly counter to all the goals that we have always attributed to welfare states.² Yet that is precisely how the vast majority of welfare states worldwide determine benefits under categorical assistance programs – programs which (in terms of sheer expenditures, anyway) can be almost twice as important as ones of general assistance.³

II

A

There are various easy ways to explain our way out of this paradox. None, however, is wholly successful. The first response is that, while

categorical assistance; thus, e.g., earnings-related supplements to the United Kingdom's basic flat-rate benefit for sickness and maternity, work injury and unemployment would, at the maximum level payable, nearly double that basic benefit.

¹ Derthick 1979, chaps. 10–14.

² Beveridge (1942, para. 302; see also paras. 9 and 304) famously recommended flat-rate social insurance schemes rather than earnings-related ones – but less for any of these reasons than on account of his obsession with self-help and his correlative supposition that people should and would secure such earnings-related supplements through "voluntary insurance" on the private market.

³ Hanson 1987.

earnings-related benefits pay more to the rich than to the poor in absolute terms, they almost invariably pay the rich proportionately less of their previous earnings than the poor. There are various mechanisms for making benefit structures progressive in these ways.⁸ However it comes about, though, such progressivity in the benefit structure would mean that even if earnings-related benefit programs pay more to the rich in absolute terms, they pay more to the poor relatively.

As it happens, this progressivity in the distribution of benefits is offset by the regressivity of the payroll taxes characteristically used to finance those benefits. Just how progressive and regressive those two are, respectively, varies between countries and is typically somewhat uncertain even within a single country. But the best guess is that the two effects just about cancel one another out in most cases.⁹

Beyond all that, though, is the larger question of principle. If the aim of welfare-state programs is to aid the poor, why should we not insist upon a program that pays more to the poor, absolutely as well as relatively? Why should we content ourselves with a program which, while less biased against the poor than the distribution of earnings itself, nonetheless pays more to the rich than the poor? There are good pragmatic reasons for such arrangements, of course. But nothing that has been said so far provides any principled justification for accepting welfare arrangements that are only half as pro-poor as they might be.¹⁰

None of the standard principled defenses of the welfare state provides an answer. Communitarian goals and egalitarian ones couched in terms of minimizing the absolute difference between rich and poor are necessarily frustrated. Relative-egalitarian and poverty-reduction goals are harder to pursue in this way. It is, for example, a point of pride among U.S. policy-makers that they have usually managed to hold the ratio of maximum to minimum benefits paid under earnings-related programs to around four to one. But if we must spend four dollars on the nonpoor in order to get one dollar to the poor, a program would have to involve truly vast expenditures overall to alle-

⁸ One is to build that progressivity explicitly into the benefit formula, as in the United States (Derthick 1979, p. 256). Another is to impose ceilings on benefits, either directly (specifying a maximum level of benefits payable) or indirectly (specifying a ceiling on how much of one's previous earnings "count" toward vesting higher benefit entitlements).

⁹ Pechman, Aaron and Taussig 1968, pp. 246-7. Derthick 1979, pp. 254-8, 290-1.

¹⁰ Diminishing marginal utility of income might be one explanation. A larger payment to those who already have more is worth the same to them, in utility terms, as a smaller payment to those who have less. But that rationale deals in terms of present means, whereas the practice ties benefits instead to past earnings. That suggests that the stabilization rationale in Section V better captures the true rationale underlying the practice.

viate poverty by very much at all. Likewise, we would have to distribute an awful lot of money at the 4:1 ratio to make much of a dent in the 10:1 ratio that marks the difference in income between the top fifth of earners and the bottom fifth."

Thus, all that the argument from the progressivity of the benefit formula succeeds in proving is that things are not as bad as they might be. Nothing in it shows that things are as good as they should be. There is, in it, no principled defense of the failure to provide more benefits to the poor than the rich in absolute as well as proportional terms under categorical assistance programs.

B

Second, there is an argument that purports to show that despite historical differences in their earnings, all recipients of categorical assistance are alike at the time they receive assistance. After all, earnings-based considerations enter into the calculation only after some categorical hurdle has been crossed; and the nature of the categorical hurdles is such as to guarantee that, however high claimants' previous earnings, they will have (largely) ceased by the time claimants qualify for assistance under any of these headings. Take the case of those qualifying for social benefits by reason of unemployment. They may have earned a lot in the past, but if unemployed then by definition they have no earned income in the present. The categories describing persons qualified for sickness, disability, maternity, or old-age benefits are similarly crafted in such a way as to disqualify anyone still in paid employment.

What this rejoinder amounts to is a suggestion that we regard the categorical tests that constitute the first cut in programs of categorical assistance as a rough-and-ready surrogate for a means test. They ensure, in effect, that anyone qualified for benefits at all will have little or no earned income in the present period. True, categorical assistance programs then go on to pay more, in earnings-related benefits, to those who had historically earned more, and that requires further explanation. But at the very least, they pay it to people who are, just at the moment, poor – in this particular sense, anyway.

That is not the only sense that matters, though. Even if reduced to zero earnings in the present period, people who have historically earned more are typically still better off in two connected respects. Generally, those who have earned more in the past will have accumulated a greater stock of capital (savings, investments, etc.) on which they can draw in the present. While some of those assets may not be

" Derthick 1979, p. 257. Atkinson 1975, p. 52.

in particularly liquid forms, in principle it should always be possible to borrow against them to tide you over. Furthermore, those who have historically earned more will typically have more unearned income (interest on savings, dividends from investments) coming in, even in periods during which they are cut off from their ordinary sources of earned income.

In practice, then, little comfort can be taken from the thought that, however much they used to earn, people who qualify for categorical assistance earn nothing now. They are still better off, in all probability, than those who have historically earned less.

C

Third is the "fairness" or "equity" rationale for making benefits earnings-related.¹² This argument frankly admits that the rich get more of those benefits but goes on to say that it is only fair that they should do so. They deserve more benefits because they paid more for them.

Categorical assistance is almost invariably organized on a "contributory" or "social insurance" basis. Some proportion of a person's paycheck is deducted and deposited into a separate trust fund. Those "contributions" serve – politically, morally and psychologically, as well as legally – to vest the person's subsequent entitlements to benefits.¹³ Deductions being linked in this way to earnings also means that those with higher earnings will have made greater contributions (at least in absolute money terms) to such schemes. In the language of insurance, the rich will have paid higher premiums, and other things being equal the larger the premiums you have paid (or the greater the contributions you have made, more generally) the more you are entitled to receive if the stipulated contingency occurs.¹⁴

Underlying the argument that benefits should be earnings-related because contributions have been is a deeper claim that people should get all that they pay for and only what they pay for. That proposition, if universally implemented, would preclude what we ordinarily regard as the characteristic function of the welfare state, that is, the

¹² Derthick 1979, chap. 10.

¹³ As Franklin D. Roosevelt explained, "We put those payroll contributions there so as to give the contributors a legal, moral and political right to collect their pensions. . . . With those taxes in there, no damn politician can ever scrap my social security program" (quoted in Derthick 1979, p. 230).

¹⁴ Chief among the "other things" that must be equal is the magnitude of the risks each party runs. But that does not explain why the rich should pay systematically higher premiums. True, they have a better chance of living longer, and hence collecting more in old-age pensions. In every other respect, though, their projected claims would presumably be less than those of the poor, who therefore should, on strict actuarial logic, be the ones paying higher premiums.

transfer of resources from rich to poor. It is, indeed, notable that programs of so-called social insurance deviate systematically from the fundamental principles of true insurance at precisely that point. The first principle of real insurance is risk rating, charging each person an insurance premium proportional to his true actuarial risk. One consequence of such a practice, however, would be that those most in need of protection, and least able to pay for it, would be charged most for it. Therefore, it is hardly surprising that the quasi-insurance components of social welfare programs adopt different rules enforcing, in effect, compulsory risk pooling. In programs of social insurance, people with very different risk profiles are forced to share the same pool, paying the same premiums for the same levels of coverage, regardless of their very different probabilities of having to file a claim. Whatever the justification for such a practice, it falls outside the logic of insurance, strictly speaking – as everyone from Hayek to Titmuss is at pains to emphasize.¹⁵

In one way, that rejoinder to the fairness rationale for earnings-related benefits simply restates the paradox with which this chapter began. If the fundamental aim of the welfare state is to supplant the market for the distribution of a limited range of goods and services, then why do some of the welfare state's very own programs incorporate precisely those same market-based criteria of people's worth by linking benefit levels to people's past earnings? If programs of social insurance should deviate from strict principles of true insurance in some respects (e.g., not making premiums reflect actuarial risks), then why should they not deviate in other similar respects (by, e.g., refusing to pay larger benefits to people just because they paid larger premiums)? All of these are just more pointed ways of asking the same basic question with which we started – if the basic ethos of the welfare state is to help the poor, then why do some of its programs pay more to people the more they have earned?

Nothing I have said so far counts as decisive counterargumentation to the fairness rationale. It does help isolate the queerness of the practice here in view, however. Why should benefit levels in categorical assistance programs be "fair," in this narrowly market-based way, when so much of the rest of the welfare state deviates so systematically from such standards of fairness?

At root, the fairness rationale for earnings-related benefits relies on notions of people's moral deserts. The basic claim is that people who have paid more deserve to get more. This desert-based logic breaks down at several points, however.¹⁶ First and most straightforwardly,

¹⁵ Hayek 1960, chap. 19. Titmuss 1968, pp. 173–87. Goodin 1988, chap. 6.

¹⁶ Goodin 1988, chap. 10.

at least some of these contributions on which people's differential entitlements are predicated were made not by the person himself but rather by someone else on his behalf. Survivor's benefits are paid not on the basis of the survivor's own social insurance contributions but, rather, on the basis of the deceased person's. In most systems of social insurance, employers are required to make payments at least as large – and often up to twice as large – as the employee's own into his social insurance account. Hence what is involved here is not, strictly speaking, a matter of paying more in benefits to those who paid more in contributions toward funding those benefits. It is instead a matter of paying more in benefits to those who either paid more themselves or had more paid on their behalf.¹⁷

In some respects, this difference might not matter much. A savings account opened in my name is just as much mine, whether the money that was deposited into it came from me or from my grandmother. If the issue were simply one of property rights, that would be the end of the matter. There is a deeper issue here, however. How can we justify vesting people with such property rights in social security entitlements, in the first place? The fairness rationale attempted to provide an answer, in terms of people's moral deserts. But for arguments cast in terms of deserts to work, the meritorious performances must be the person's own. Desert is predicated on the basis of *his* character or *his* past performances. A person can deserve something, therefore, only if he did something meritorious himself. For purposes of moral deserts, it is simply not enough to have had something meritorious done on your behalf.

The desert-based argument breaks down at a second point, too. In order for desert claims to be predicated upon a person's past performances, they must have been voluntary performances. Just as people deserve no blame for coerced or otherwise involuntary actions, so too do they deserve no credit in such situations, either. Now, contributions to social insurance schemes are coerced in just this way. Indeed, this marks out the second important point of distinction between social insurance and private insurance schemes. Under private insurance schemes, participation is ordinarily completely voluntary. Under social insurance schemes, participation is ordinarily compulsory for everyone who falls within the ambit of the scheme; and contributions to the scheme are extracted under threat of penal sanction (as their

¹⁷ The employer's contribution might be considered a fringe benefit paid to the employee in lieu of higher wages. (Apparently it is explicitly so in, e.g., German wage bargaining.) Then that might count as the employee making a sort of indirect deposit to his own account. But having no option of demanding the wages instead, he "accepts" that contribution in lieu in such an attenuated sense that the second objection, to be discussed, remains.

popular designation – “social security taxes” – makes plain). Such coerced contributions can form no basis for a desert claim. For purposes of moral deserts, it is simply not enough to have been forced to do something meritorious.

The fairness rationale for earnings-related benefits, therefore, is not only queer but doubly flawed. The argument that people with higher earnings deserve higher benefits because they made higher contributions presupposes (1) that they made those contributions themselves, rather than having had others make them on their behalves, and (2) that those contributions were made voluntarily. Both assumptions are substantially in error, the latter one virtually completely so.

III

None of those arguments suffices to justify linking social welfare benefits to previous earnings in such a way as to ensure that more benefits go to the rich than to the poor. For a more satisfactory explanation-cum-justification of such practices, we need to reconceptualize somewhat the function of the welfare state.

At least in part, the welfare state is not really aimed at aiding the poor or even at aiding people when they are poor. It does that, too, of course, principally through its general assistance components. But there is a further, independent, and really quite important function that it also serves – principally through its categorical assistance components. That is to provide a certain measure of stability to people’s economic affairs. At least in part, the function of the welfare state is to underwrite, and in that way to help stabilize, people’s market-based earnings expectations. Its job is, first, to smooth out the peaks and troughs in their earnings patterns, bringing their short-term rewards more into line with their long-term average earnings. And when through some unanticipatable event their long-term earnings expectation suffers a sharp and irreversible decline, the role of the welfare state is, second, to ease the transition from the old, higher expectation to the new, lower one.

This analysis is far from novel.¹⁸ Indeed, some such intention is signaled in the very names of the programs. They are called, variously, “income maintenance” or “social security” programs, after all. Commentators who are fixated on the question “What does it do for the poor?” have their own interpretation of those phrases, of course.

¹⁸ Earnings replacement while incapacitated was clearly the aim of early workmen’s compensation schemes (Nonet 1969, pp. 21ff.). The point of social insurance more generally is to make “provision against interruption and loss of earning power” (Beveridge 1942, para. 12; see similarly Pechman et al. 1968, p. 55).

For them, it is all a matter of maintaining people's incomes above a floor set by the poverty threshold; the "security" involved, in their view, is securing people against impoverishment. There is no need for this very narrow interpretation, though. Here I shall show that these programs do, and are right to do, precisely what their titles suggest: provide stabilization of a more general sort.

One component of the welfare state – the contributory old-age pension – is already widely understood to serve roughly the function I here attribute to a whole broad range of such programs. It is commonplace to say that such pensions work to help equalize income over a person's lifetime, taxing him in his high-earning years to pay for income supplements for him in his nonearning years. Again, those fixated on the question "What does it do for the poor?" tend to suppose that the justification for that practice falls straightforwardly out of the justification for redistributing resources from rich to poor more generally. Others, however, appreciate that this redistribution – though undeniably compulsory – is more justifiable, somehow, for its being intrapersonal. To explain why, I suggest that we need to appeal to the values of stability and security in a person's life.

Much in the administrative detail of earnings-related benefit programs also supports the sort of interpretation that I am here advocating. Consider, first, the question of why some categorical assistance benefits are earnings-related, while others are not.¹⁹ Virtually all are, as already established, but there is one conspicuous exception. "Family assistance" or "child benefit," uniquely among categorical assistance benefits, is paid (invariably, in all nations of the world with such a program at all) without respect to past earnings of the mother, head of household, or anyone else.²⁰ The reason, I submit, is that uniquely among categorical assistance payments these were never meant as replacements for lost earnings.²¹ This pattern is wholly in line with my interpretation of earnings-related benefits as devices for stabilizing personal fortunes. In those terms, too, it would be important that benefits be related to earnings where benefits take the place of lost earnings, but there would be no need for them to be where they do not.

Consider, second, the way in which claims for categorical assistance

¹⁹ Of course, as already discussed, programs of means-tested general assistance are organized around their principles altogether; and given those principles, it is only natural that their benefits should be related to recipients' needs rather than to their past earnings. So, in looking for genuine anomalies, we must confine our attention to programs of categorical assistance. Anomalies can be found even there, however.

²⁰ U.S. DHEW 1978, col. 5, row 5, for each country entry.

²¹ Notably, "maternity benefits" almost invariably are earnings related, and of course they do substitute for lost earnings of the mother immediately around the time of delivery.

must be made promptly if they are to be honored at all. British rules seem fairly typical in this regard. They stipulate that a person will not ordinarily be paid benefits for a day's unemployment unless he claims them on the very day; he will not ordinarily be paid sickness, invalidity, or injury benefits for a claim more than twenty-one days past, or for disability, retirement, or maternity hospitalization more than three months past. Most of those time limits are extendable in special circumstances, but all are subject to an absolute time limit of twelve months.²²

Such strict time limits for claiming make these odd sorts of entitlements. True, one's claim to property of a more ordinary sort might likewise lapse with time, but in land law it characteristically takes something like a dozen years of adverse possession before that happens.²³ The real reason for these social security rules may well be baldly pragmatic, designed to help hard-pressed bureaucracies under severe budgetary constraints trim their rolls.²⁴ In principled terms, though, the practice seems something of a mystery.

The key to resolving this mystery, I suggest, once again lies in the role of categorical assistance in stabilizing people's fortunes by replacing lost earnings and thereby avoiding unnecessary disruptions to their lives. A hard man might say that if people did not claim the benefit promptly then presumably they did not really need it (they got by on savings, or some such). That is not necessarily true, of course. Many who need the benefits most are least well informed of their entitlements and fail to claim them on time in consequence. What is necessarily true of very late claims, though, is this. By the time they are made, people will already have suffered precisely the sorts of disruption that they were designed to protect against. That rationale for paying the benefits is simply no longer relevant.

Consider, third, the way in which categorical assistance for the relief of temporary distress is for a fixed period paid at a higher rate than long-term assistance. Details vary from country to country, naturally. But the typical pattern, across OECD and Comecon countries alike, is this. The invalidity pension, paid to the long-term sick and disabled, is usually paid at the same rate, as a percentage of previous

²² Partington 1978, pp. 24-32.

²³ Buen (1982, chap. 27), summarizing the U.K. Limitation Act of 1980. Similarly, tort claims are subject to a statute of limitations, but again it runs for much longer and, in any case, is rationalized essentially on grounds that evidence is unreliable after a protracted period (St. Leonards 1852). Such rationales hardly apply to the case of social security claims of the sort here in view. Reliable evidence of hospitalization is easily obtained, even after several months have passed, for example.

²⁴ Lipsky 1984.

earnings, as the old-age pension. But short-term sickness and maternity benefits are usually paid at rates at least half again higher, for periods of between six months and a year, after which claimants are shifted over to invalidity pensions.⁴³

Why temporary distress should have such a greater claim on our sympathies is, again, something of a mystery. Presumably those who are destined to suffer the same misfortune forever would be worse off and have a greater claim in consequence. Again, there are perfectly good pragmatic reasons for treating them worse – short-term generosity is cheap; long-term generosity gets expensive. But if we are to explain the practice in principled terms, once again we must have recourse to the stabilization of expectations rationale.

In those terms, the function of welfare benefits is to tide people over, if their distress is temporary, and to ease the transition from higher expectations to lower ones, if their plight is to be permanent. Generous short-term assistance – sometimes amounting to 90 or 100 percent of previous earnings – serves both goals. But if, after six months to a year, the situation has not changed then it is time for people to start coming to terms with their new circumstances. They will have had adequate warning, and more than enough transitional assistance, to make the requisite adjustments.

IV

The need for some such stabilization scheme is evident from recent studies of the way in which a surprisingly large fraction of the population's economic fortunes fluctuate wildly from one year to the next.⁴⁴ In aggregate, patterns of income distribution seem remarkably stable over time. The top decile of earners gets roughly the same portion of national income from one year to the next and, indeed, one decade to the next. But when the University of Michigan Panel Study of Income Dynamics followed the economic fortunes of five thousand

⁴³ U.S. DHEW 1978, comparing row 2, col. 5, with row 1, col. 6, of each country's entry. This is not just an artifact of comparing across program categories, however; sometimes (as in the two Germanies and Austria, e.g.) sickness benefit itself is paid at a higher rate for the first half-dozen or so weeks, and then at a lower rate thereafter. Broadly the same pattern holds if you compare the temporary and permanent disability pension provisions (row 3, cols. 4 and 5), although there you do find the very occasional anomaly (such as France, which pays more for permanent than temporary disability pensions).

⁴⁴ Presumably the reason we worry more about stabilizing income than capital is precisely that, for most people most of the time, capital stocks do not fluctuate nearly so widely as income streams. That, together with the fact that capital really matters to people's lives only when transformed into income, explains why the welfare state has income stabilization but not capital stabilization schemes.

American families over the period of 1968–78, it found that “this apparent stability” at the level of gross aggregates “is an illusion produced by the offsetting effect of many substantial upward and downward changes” at the level of individual families. “Fewer than one-half of the population remained in the same economic position from the late 1960s to the late 1970s, while one-third had dramatic improvements in their economic well-being and one-fifth had dramatic declines.”²⁷

This volatility manifests itself most generally and in many ways most importantly in dramatic fluctuations in family earnings from one year to the next. The effects of the family’s second earner’s joining or leaving the work force are, surprisingly, swamped by the effects of variation in the hours worked (and hence wages earned) by the primary earner “owing primarily to variations in overtime, second jobs or short periods of unemployment.”²⁸ Even for white men, who are presumably the most stable group of earners in the sample, the average year-to-year change in annual income was 25 percent; 40 percent saw their annual income fluctuate by 10 percent or more six or more times in the decade; “no individual completely escaped a decline in earnings, and almost 60 percent had declines in at least four years.”²⁹

The earnings patterns of females and of nonwhites is still more volatile, of course. But the striking conclusion of these findings is that, even among white males, “no identifiable group – not the more educated, not union members, not even higher-income persons – seems to be immune from these changes in year-to-year income. There is no evidence that there are secure, protected niches in the economy. . . . Variability rather than stability and regularity characterizes the working lives of most men.”³⁰

This common experience of fluctuating earnings in general is matched, and to some extent caused, by common experience of the particular sorts of misfortune that welfare state programs are designed to cushion against. The Panel Study of Income Dynamics, again, found that “major unemployment” struck 29 percent of married men at least once during the eleven years, 1968–78. Involuntary job changes affected 22 percent. Work loss due to illness affected

²⁷ G. Duncan 1984, p. 3.

²⁸ G. Duncan 1984, pp. 95, 98, 100. The average variation for all white male heads of households over the ten years is 15 percent. For three-quarters of them the variation exceeded 20 percent in at least one year; for a quarter of them, this pattern was so frequent that their ten-year average rose above that mark.

²⁹ G. Duncan 1984, p. 121.

³⁰ G. Duncan 1984, p. 119.

28 percent. Disability struck 30 percent. The first two affected predominantly younger (and to a lesser extent middle-aged) men; the third, predominantly middle-aged (and to a lesser extent younger) men; the fourth, predominantly men over sixty, but to a lesser extent middle-aged men as well.¹¹ One way or another, though, every age group is at substantial risk of needing categorical assistance from the welfare state to protect against fluctuating fortunes in the labor market.

Significantly, however, the fluctuations seem to work in both directions for the vast majority of people. They may need welfare to tide them over a bad patch, but they are soon back on their feet again. This is evident from the pattern of welfare utilization over time. The Panel Study of Income Dynamics reports that between 1969 and 1978 "one out of every four Americans lived in a household that received income from one of the major welfare programs at least once," but fully "one-half of the persons who lived in families where welfare benefits were received once in a decade did not receive it in more than two of the ten years."¹² Thus, it seems that most welfare recipients are not long-term "welfare dependents" in any important sense – only about 2 percent of the sample seem to have depended upon welfare benefits for more than half their income for eight or more of the ten years.¹³ Most people, instead, seem to have relied upon welfare benefits as a form of transitional assistance.

The conclusion to be drawn from all these statistics is aptly summarized by the Panel Study of Income Dynamics report itself:

No broad demographic group in our society appears immune from shocks to their usual standard of living, shocks resulting from rapidly changing economic or personal conditions. For men, the shock often comes in the form of an involuntary job loss; for married women, divorce or the death of the spouse is often the precipitating event. Such events may not always be totally unavoidable, but few people are immune to occasional economic misfortune, and when it strikes, welfare serves as a kind of insurance for them, providing temporary assistance until they are able to regain their more customary levels of living. . . . These people are "digging out" following a disaster. Welfare assists them during that process and then, in time, is left behind.¹⁴

¹¹ G. Duncan 1984, p. 27.

¹² G. Duncan 1984, p. 90.

¹³ G. Duncan 1984, p. 75.

¹⁴ G. Duncan 1984, p. 90. Although Duncan's findings pertain only to the United States, strictly speaking the precipitating events here mentioned – involuntary job loss, incapacitating accident or illness, divorce or death of a spouse – are all equally common events in other developed societies. Hence there is some reason to suppose that the same findings would emerge there as well.

So far, I have shown that the earnings-related benefits of the welfare state's categorical assistance programs are essentially devices to stabilize people's fluctuating economic fortunes. That, apparently, is what they were designed to do. That, apparently, is the way in which most recipients utilize them. I have left until last questions concerning the moral justifiability of a public policy of underwriting people's market-based economic expectations in such ways. It is to this topic that I now turn.

The first thing to be said, I suppose, is that the recipients of such assistance are themselves better off in consequence of such policies. Bouts of incapacitation or unemployment are disruptive, in all sorts of ways. But earnings-related categorical assistance programs help to ensure that temporary exigencies do not force any unnecessarily dramatic alterations in the course of a person's life. So long as welfare benefits are sufficiently large really to replace most of his lost earnings – and such temporary forms of assistance (for sickness and unemployment, e.g.) typically are, often amounting to 90 or even 100 percent of previous earnings – a person will be able to resume his life, economically at least, much as before once his misfortune has passed.

Those who are in actual receipt of benefits are not the only ones to benefit in such ways from such programs, though. Everyone who was *ex ante* at similar risk of similar misfortune would benefit from the insurance that such programs provide. Even if they never have to claim benefits, at least they know that they can count on such bridging assistance should they ever need it. That knowledge allows them to proceed with their plans in confidence, knowing that it will take more than merely passing misfortune of the sort that might well befall any of us to spoil their plans. How many people are in a position to require such reassurance is, of course, an empirical question, the answer to which naturally varies from country to country. On the evidence of Section IV, it seems that in the United States anyway most people are in broadly the same boat in this respect. Fully a quarter of Americans benefited from welfare at some point in the decade, and inspection of the personal characteristics of those who did suggests that no demographically identifiable group is immune to, or even at substantially less risk from, such passing misfortunes.¹⁵

¹⁵ The primary exception would seem to be black Americans, who undeniably do suffer substantially greater risks of lasting misfortune. In terms of the logic of this chapter, perhaps they should therefore have lower expectations-based welfare entitlements. But that only serves to emphasize the importance of supplementing expectation-based principles of categorical assistance with the egal-

Thus, everyone – recipients, and everyone else qua potential recipient – benefits from social welfare policies stabilizing incomes in the face of temporary setbacks that people may suffer. That does not quite settle things morally, however. Everyone benefits, but some benefit more than others. Some people's incomes are stabilized at a higher level than others', and more money is spent in so doing. That seems morally problematic, in a way that simply "everyone benefiting" does not.

There are various ways to try to make that outcome seem less problematic. One particularly bold strategy is this. Suppose we have got the primary income distribution "right," in the sense that we have good grounds for supposing that those who do earn more should earn more. What those grounds might be can safely be left open, here. Maybe it is a perfectly planned economy, in which all wages are set at just the right level; maybe it is a perfectly competitive market economy, in which all wages have reached the equilibrium levels. No matter. The next step in the argument is to assert that once we have the primary distribution right, the secondary distribution should mimic it. That amounts to saying that whatever it is about the level of income that a person ordinarily secures through earnings that makes that the "right" level also makes it right that that level be reproduced through welfare benefits, should his earnings temporarily cease.

There is no particular reason to suppose that that should be so, however. Certainly neither of the accounts alluded to earlier of what makes people deserve what they earn when in work would imply that they deserve the same when out of work. Socialists would justify higher earnings by virtue of a worker's greater social contribution. But when he ceases working, for whatever reason, that greater contribution ceases – and with it his claim to higher income. Similarly, the market rewards scarce skills only insofar as they are actually employed.

Other less bold strategies are of course available. A singularly economic approach, for example, is to argue that the reason we relate earnings to benefits paid people when they are involuntarily out of work is to provide incentives for them to maximize earnings while in work. This analysis seems rather far-fetched, though. Surely for most people, the impetus to maximize earnings comes principally from a desire for the earnings themselves, rather than from any reflection upon the (typically uncertain, and anyway much weaker) effects of those earnings on social welfare entitlements.*

itarian, poverty-reduction rationales underlying programs of general assistance, alluded to at the outset of this chapter.

* Except perhaps among marginal groups in the work force, such as those on the verge of retirement or a family's second income-earner (Hammermesh 1979).

Another superficially more plausible approach would be this. Presumably it is desirable, in terms of utilitarian moralities and perhaps others as well, to help each person achieve the highest standard of living that he can sustain on a long-term basis. Insofar as short-run fluctuations threaten to undermine people's long-term attainments, then a goal of helping people to maximize the latter would commit us to providing a social cushion against the former. And insofar as any large deviation from a person's long-term mean performance would undermine his overall attainments, such assistance must be proportional to his long-term expectations. That, too, would rationalize earnings-related benefits.

There is a valuable insight embedded in that argument, which I shall put to rather different use shortly. But cast in utilitarian terms of welfare maximization, the argument will not go through. Were we dealing with goods of "ultimate" or "end-use" value, then the way to maximize aggregate social welfare would indeed be to maximize each person's welfare independently of any other's. But in the welfare state it is programs of general rather than categorical assistance that dispense goods of end-use value (food, clothing, shelter, etc.) through in-kind assistance; and those programs, recall, respond to sheer need and are indifferent to past earnings. Benefits are earnings-related only with respect to categorical assistance, where the benefits dispensed are invariably in the form of cash payments. Now, cash is a paradigmatically "instrumental" good, valuable merely as a means to obtaining other things that are of value but not in itself of any value. Where the goods involved are merely of instrumental value, there is no reason to suppose that aggregate social welfare would be maximized by maximizing each person's holdings. On the contrary, given certain assumptions about the competitiveness of the relative power conferred by instrumental resources over the allocation of a fixed stock of end-use goods, there is reason to believe that aggregate social welfare would be maximized by equalizing rather than maximizing holdings of instrumental goods like cash.¹⁷ If anything, then, the welfare-maximization argument provides yet another ground for opposing rather than favoring earnings-related benefits stabilizing some people at higher incomes than others.

For a really satisfactory justification of that policy, we must look elsewhere. The most satisfactory one, in my view, is that some morally important outcomes cannot be realized unless people can count on their being realized. This is a point usually pressed by rule utilitarians against their act utilitarian rivals. I can plan my own activities in a

¹⁷ See Chapter 15 in the present volume (building on Hirsch 1976; Sen 1983; and Goodin 1988, chap. 9).

maximally beneficial way only if I can rely upon you to keep your promise, come what may, without act utilitarian calculations of consequences in each particular case.³⁸ But the point is really a much more general one and admits of a more personalistic formulation as well. Where rule utilitarians would focus upon the benefits to society at large from people's being able to form reliable expectations about the future, others (such as Bernard Williams) would focus upon the value of coherence and continuity within my own life achieved only by having reliable expectations about the future.³⁹

The crux of the matter, for both personalists and rule utilitarians, is that the plans and projects that constitute people's lives are complex things, requiring many interlocking parts to be fitted together. Often those contributions must (sometimes contingently, sometimes necessarily) come sequentially rather than simultaneously. Where that is so, earlier contributions would have been wasted if later ones are not forthcoming; and given that, people might not risk making those earlier contributions at all, in the absence of some good guarantee that later ones will indeed be forthcoming. Whether we are utilitarians counting on other people's subsequent contributions to our interdependent projects, or whether we are personalists counting on being able to sustain our own ongoing projects through time, it is valuable for us to have some guarantee that the fluctuating fortunes of ourselves and others will not needlessly interrupt our long-term plans.⁴⁰

Thus, there is a certain value – both to individuals and to society – from people's being able to enter into long-term commitments. Insofar as those commitments require rough stability of a person's earnings over time, that value will be well served by social welfare programs paying people benefits proportional to their past earnings, at least in cases of temporary interruptions to their ordinary income stream.

It is an open question whether we should, on similar grounds, also sustain the previous earnings of those who suffer more permanent disruptions to their income (through disabling accidents, debilitating illnesses, or long-term unemployment). Certainly such open-ended expenditure commitments are costly, and societies may well shy away from them for that reason alone. But there are matters of principle at

³⁸ Hume 1739, bk. 3, pt. 2, sec. 2; 1777, sec. 3, pt. 2. Hodgson 1967. Harsanyi 1977b.

³⁹ Williams 1973a; 1981. Wollheim 1984. Nozick 1981.

⁴⁰ "Needlessly" in deference to the caveat entered later about the inevitability, and hence desirability, of readjusting life plans in light of permanently changed circumstances (see further Chapter 10). Note that even the personalist, concerned only with one's accomplishment of one's own plans and projects, ought worry about fluctuations in the fortunes of others insofar as those would prevent them from doing something that they must do if one's own plans are to succeed.

work here, too. In one sense, we should sustain the previous earnings of permanent welfare recipients for the same reasons we should those of temporary ones. Both have long-term plans predicated on the presumption of a continuing resource stream of that sort. In another, more important sense, however, those who have been permanently incapacitated would be living a lie to persist with their same plans, just as before, as if nothing had happened to them. Of course we can artificially sustain their endeavors through generous social assistance for as long as we like. But that very artificiality deprives the outcome of any status as "their" accomplishment, just as surely as if a cripple were to participate vicariously in a marathon through the running of her daughter.

Reconciling oneself to one's fate is not just another virtue, the value of which is to be set off somehow against the value of framing plans for oneself. Instead, we must conclude that whatever value we see in people's framing plans and projects for themselves is conditional upon their being realistic plans and projects, appropriate to the person's circumstances.⁴ After a time, the permanently incapacitated ought to be expected to readjust their plans and projects to their new circumstances.

Still, categorical assistance payments designed to replace lost earnings ought to be proportional to – and, indeed, a large proportion of – past earnings, at least in the short term, even for the permanently incapacitated. After all, it takes time to see that some states are truly permanent. Beyond that, it takes time for people to discharge or renegotiate previous commitments made on the basis of prior expectations. If we think that there is value in people's entering into long-term commitments of this sort, then we must be prepared to provide at least transitional assistance of a powerfully earnings-related form while they extricate themselves from them, when changed circumstances force them to do so.

Furthermore, some of those long-term commitments can never be completely discharged or renegotiated. To some extent, the plans one made before disaster struck, and predicated upon one's older and rosier expectations, will always be with one. To help people discharge those obligations *ex post* of the disaster – and to reassure people *ex ante* of any disaster that it is safe to give and to rely upon such commitments – we ought to provide welfare benefits of at least a weakly earnings-related form to the permanently incapacitated, even in the long term. That assistance need not constitute such a large proportion

⁴ Furthermore, unpleasant surprises – often of a sort for which cash transfers constitute necessarily inadequate compensation – are, realistically, a fact of life. People ought therefore also to be expected to frame plans in such a way as to be minimally sensitive to anticipatable buffeting.

of their previous earnings, on the assumption that they can divest themselves of at least some of their previous commitments. But it ought nonetheless be proportional to previous earnings, on the assumptions that at least some of those previous commitments cannot be extinguished and that the commitments entered into by those who used to be rich will be more expensive to honor than those entered into by those who used to be poor.⁴²

The value of promoting security and stability in people's lives is only one value among many that we would like our social arrangements to serve. Pursuing that goal through earnings-related social welfare benefits necessarily conflicts, in ways sketched at the outset, with goals of equality, of community, and perhaps even of welfare maximization. Nothing I have said here should be taken to imply that the values of security and stability either override or even weigh substantially more heavily than those others. In cases of conflict, those other considerations may well, on balance, prove decisive. Far from denying that that might happen, I have here been merely concerned to show that there is something (albeit nothing decisive) to be said in defense of programs of earnings-related benefits designed to stabilize people's expectations. That in itself is something of an accomplishment, though, given the original mystery with which I began.

VI

The claim of the preceding section is just that security and stability are morally worthy goals and that they are well served by earnings-related social welfare benefits. But to say that those goals are well served by such programs is not to say that they are "uniquely," or even "best," served by such programs. To justify why earnings-related social insurance should be both compulsory and state-provided, we must show why such a scheme is superior to voluntary private insurance on the one hand and to compulsory private insurance on the other.

Fortunately, there is a large literature within economics on the failure of insurance markets, much of which has been shown to apply

⁴² The assumption of lingering income-related expenditure commitments, while probably true in general, may well vary greatly across individuals. Claus Offe (personal communication 1994) suggests that we pick this up directly, rather than indirectly, by subjecting earnings-related social benefits to heavy progressive taxation, for which there are generous exemptions for life plan-specific expenditures. The trick there would be in specifying those exemptions with any degree of generality. My own hunch is that under this strategy either too little (or, over time, too few of the right things) would gain exemption or else officials would have to be granted too much discretion in doing so (see Chapter 14 in the present volume and Goodin 1988, chap. 7).

with full force to the sorts of hazards involved in welfare state services." Some of the arguments have to do with notions of "moral hazard" (the danger that people, once insured, have less incentive to take care) and of "economies of scale." Others have to do with risks being interdependent rather than independent (given the interconnections within the economy, the probability of a steel worker's being unemployed is not independent of that of a car worker's, or even of a garment worker's). Still others have to do with the impossibility of the relevant actuarial calculations (no one, setting up a retirement annuity in 1945, could have guessed what his life expectancy or economic requirements would have turned out to be when retiring in 1990). All those arguments justify making social insurance compulsory and making the state the underwriter of last resort of such schemes.

Two such arguments are of particular relevance to the themes of this chapter, though. One has to do with what economists dub "externalities." As the rule utilitarian argument canvassed in Section V suggests, my life plans are to some extent intertwined with yours. If your plans are upset by some interruption in your earnings, there will be some serious consequences for mine as well. Most dramatically, if you go bankrupt, your own plans are certainly interrupted – but so too are those of your creditors. Now, you might be tempted to insure yourself against the consequences of your interrupted earnings for your own plans; but you would have little motive to buy insurance to cover me for the consequences that your misfortune might have for me. The externality argument thus justifies compulsory insurance of earnings on essentially the same grounds as we would justify compulsory inoculation against contagious diseases.⁴¹

There is another reason justifying us in making it compulsory for you to be insured against catastrophes that would seriously interrupt your earnings, quite apart from that, however. This second reason has to do with the fact that we have good grounds for regarding your ex ante risk preferences – those that would be manifested in any voluntary insurance decision – as highly suspect. What matters morally, of course, is how you would feel after the fact, once the catastrophe has occurred. As Sen says, "We are interested in tomorrow's satisfaction as such, not in today's assessment of tomorrow's satisfaction."⁴²

⁴¹ Arrow 1963. Barr 1987, chaps. 5, 8 and 9; 1989. Goodin 1988, chap. 6.

⁴² Strictly speaking, that justifies compulsory insurance against losses you suffer when someone else's earnings are interrupted, rather than compulsory insurance against your own earnings being interrupted. But as with the smallpox case, it is typically more efficient to protect yourself from the consequences of others' misfortunes by preventing those misfortunes from befalling them rather than by trying to shield yourself from knock-on consequences when they do.

⁴³ Sen 1957, p. 746; see further Goodin 1982b, chap. 3.

Yet there is abundant evidence in recent psychometric studies for supposing that people both underestimate the probabilities of catastrophes happening to them and underestimate how awful they would feel if they did.⁴⁶ Insofar as *ex ante* preferences really are systematically different from *ex post* ones, relying on voluntary private insurance in this area would make the wrong consumer sovereign.

VII

It is standardly said, in objection to all expectation-based models of social obligation, that something is required to underwrite the reasonableness of the expectations if they are to generate obligations at all. Unreasonable expectations, even if relied upon, can ground no claims.⁴⁷

But those who plan their lives in the expectation that their present earnings pattern will continue unabated are not being so obviously unreasonable, though. In good Bayesian fashion, they are letting past experience shape their expectations about the future. What could be more reasonable than that? Even once misfortune has struck them, it is not necessarily unreasonable for them to expect that they will resume their former lives again in fairly short order. After all, most people do.

On the foregoing evidence, however, what people ought to expect (from the past experience of themselves and their fellows) is a pattern of fluctuating economic fortunes. Fully a quarter of the population suffered setbacks so severe as to force them onto welfare rolls sometime during the 1980s; no subgroup of the population is substantially more immune than any other; hence, presumably everyone must know someone (or someone who knows someone) who has recently suffered a serious setback. Knowing that disaster might strike, people ought reasonably to plan against it – by savings and private insurance to replace lost earnings where possible, or by scaling down their plans so they do not depend upon such high earnings otherwise.⁴⁸ If people plan on the assumption that their earnings will con-

⁴⁶ Kahneman, Slovic and Tversky 1982. Goodin 1982b, chap. 8. See further Chapter 8 in this volume.

⁴⁷ Fried 1981, p. 10; cf. Goodin 1985c, p. 47.

⁴⁸ All the more, they must plan against the virtual certainties of life: growing old, quitting work, dying. In cases of anticipatable drops in earnings (such as are associated with compulsory retirement, e.g.) there is much less of a case to be made in the terms sketched here for making long-term assistance even weakly earnings-related. The most we might say is that there are some worthwhile projects which can be pursued over the course of one's working life only on the assumption that they will be pursued beyond it as well; and to encourage people

tinue at their peak levels without interruption into the indefinite future, in light of the clear evidence that should be known to them that they will not, they are indeed behaving unreasonably. Certainly they cannot appeal to the Reverend Bayes for support, at least.

What makes it reasonable for people to proceed in this way is, of course, the knowledge that the welfare state will rush to their aid if required. To say that is to make earnings-related benefits justified by a kind of self-fulfilling prophecy. People's previous earnings should be sustained, because they expect that they will be; but the only particularly good grounds that they have for expecting that to happen is the reassurance that the state will provide earnings-related benefits if the market ceases to provide earnings. Obviously, we cannot justify welfare programs by reference to expectations that they themselves induce. If when eliminating the program the expectations would disappear with it, then the existence of those expectations cannot be what justifies having the program. The justification would then swallow its own tail.

In the end, I think we are forced to say simply that a life of fluctuating fortunes is less good – for the person living it, and for others around him whose fates are linked to his – than one characterized by more stability. No particularly fancy story is required to sustain this claim, though. Someone who has to hedge against more possibilities, and keep more options open, will have a life that is internally less purposeful and coherent and externally less predictable and dependable. Such a life is less satisfying to the person leading it and less useful to others depending upon him. It is that, I submit, that makes it reasonable for people to expect the state to step in to provide stability to people's lives. It is that that justifies the state in paying welfare benefits proportional to past earnings, at least for a while.⁴⁴

to pursue those projects we will have to guarantee them the (earnings-related) wherewithal to continue pursuing them, even after earnings have ceased.

⁴⁴ Earlier versions of this chapter were read at All Souls' College, Oxford, the University of Arizona and the London School of Economics. I am grateful for comments, then and later, from Tony Atkinson, Nick Barr, Brian Barry, Tim Bessley, Chris Bertram, Tom Campbell, Howard Glennester, Julian Le Grand, Ted Marmor, David Miller, Ned Muller and Claus Offe.

Compensation and redistribution

Compensatory justice is profoundly conservative, standardly serving to restore some status quo ante.¹ The essential aim of redistributive justice, in contrast, is to alter those antecedent distributions which compensatory justice is at such pains to re-create. The two notions thus seem unalterably at odds, compensation striving to preserve what redistribution seeks to change and redistribution altering what compensation strives to preserve.

If forced to choose, rarely will it be compensatory justice that we choose to abandon. We are not prepared to let wrongs go unrighted, merely on the ground that the wronged are far richer than the wrongdoers;² we are not prepared to deny workers compensation for accidental injuries at work, merely because they were being overpaid anyway.³ Principles of compensatory justice in this way have us

¹ That is variously characterized: as the position people were in before others wronged them (compensatory damages, in the law of torts); as the position people were in before public takings of their private property (just compensation, in the law of eminent domain [Michelman 1967; Epstein 1985]); as the position people were in before changes in public policy put them out of work (compensation provisions in legislation liberalizing trade or deregulating airlines [Goldfarb 1980; Cordes and Goldfarb 1983] or extending the boundaries of national parks); as the position people were in before an accident or injury or other misfortune befell them (workmen's compensation, unemployment compensation, accidental injuries compensation, criminal injuries compensation). For comprehensive surveys, see Atiyah and Cane 1987 and D. Harris et al. 1984.

² Even the doctrine of "deep pockets" in torts is used only to decide which among multiple tortfeasors should bear the costs. It would never dictate leaving a wrong unrighted merely on the ground that the victim, though wholly blameless, nonetheless has the deepest pockets in town.

³ Unless, perhaps, that better pay incorporated a "risk premium" reflecting the risk of such accidents – in which case it might be argued that the injured worker has already been compensated once, and to compensate him again would constitute double-dipping (cf. Broome 1978b; Goodin 1982b, chap. 8).

firmly in their hold, and insofar as they do, principles of redistributive justice (seen as their polar opposites) have trouble getting any grip at all.

Contemporary contractarians of a right-wing cast make much of that point. For them, the permissibility of interventions is predicated upon the agreement (real or hypothetical) of everyone affected; and people's agreement in turn is predicated on compensation (direct or indirect, explicit or implicit) for any losses. From that pair of propositions, writers such as Nozick and Epstein conclude that any genuinely redistributive interventions must be decisively blocked.⁴ Those antiredistributivist conclusions follow, however, not from any positive arguments (which are no more than arguments in favor of compensation) but merely from the further supposition that redistribution is necessarily the antithesis of compensation.

The burden of this chapter is to show that compensation and redistribution are not necessarily incompatible. Everything depends on how you set about justifying compensation in the first place. Some arguments would indeed make compensatory justice an implacable foe of redistributive justice, but those turn out to be the least defensible rationales for compensation anyway. The rationale for compensation that is most defensible proves to be broadly compatible with a certain measure of redistribution, practiced in a certain way; it may even demand a certain measure of redistribution.

Here I shall be making a weak claim strongly and a strong claim weakly. The former – my major thesis – is a compatibilist thesis. Properly conceived, compensation is perfectly compatible with redistribution of a certain sort. More formally, the moral goal that is served by compensation does not imply that we must not redistribute at all. Establishing that involves a two-step argument, first showing what the true goal of compensation is (Section I) and then showing that that goal is perfectly consistent with redistribution of a certain sort (Section II). The latter – my minor theme, offered in a more speculative spirit – is an entailment thesis, suggesting that perhaps compensation actually entails or is entailed by redistribution in certain respects. I offer some thoughts along these lines in Section III. The major purposes of this chapter will, however, have been served by the compatibilist thesis alone. The principal aim is to block one apparently powerful argument against redistribution. If that blocking move also creates an opening for further advancing the positive case for redistribution, that is merely a bonus.

⁴ Nozick 1974, chap. 7. Epstein 1985, esp. part II.

I. RATIONALES FOR COMPENSATION

Compensation, as characterized in Chapter 11, aims to provide the "full and perfect equivalent" of what was lost and thereby to restore completely the status quo ante.⁵ There are three sorts of reasons for wanting to do so. The first has to do with the substantive rightness of the status quo ante to be restored. The second has to do with the wrongness of the process by which that status quo ante was upset. The third has to do with something more formal about the status quo ante, wholly apart from its content or the process by which it was disturbed. I shall explore each possibility in turn. To foreshadow, by arguing against the first (substantively based) rationale for restoring the status quo ante, I will eliminate that rationale for compensation which would indeed set it necessarily in opposition to redistribution; and that clears the ground, in turn, for the compatibilist thesis to be developed in Section II.

A. Restoring right outcomes

The first and most natural reason for insisting on compensation is to restore a status quo ante that embodied a distribution that was, in some sense, *substantively right*. On this hypothesis, the rightness of that distribution wholly explains why we should insist upon compensation to restore it.

Following Nozick, we can further decompose claims about the rightness of the distribution being restored. One class of claims traces its rightness to some special features of the *pattern* or *end-state* embodied in that distribution. A second class of claims traces its rightness to some special features of the *history* or *process* by which it came about.⁶ Typical of the former class are claims of egalitarians, who maintain that a distribution is just insofar as it displays a pattern of equal holdings. Typical of the latter are claims of so-called entitlement theorists, who maintain that a distribution of property is just so long as the history of its creation conforms to certain rules of justice in original acquisition and subsequent transfer.⁷

Representing as they do structure and process arguments respectively, end-state and historical analyses seem to exhaust all the logically possible ways of arguing for the rightness of distributions. Neither form of argument, however, is capable of rationalizing the

⁵ Brewer 1893, p. 326.

⁶ Nozick 1974, chap. 7.

⁷ Nozick 1974; Epstein 1985.

practice of compensation as we know it.⁶ Those arguments only privilege *certain* sorts of distributions – those displaying certain characteristics picked out by each kind of argument. Compensation, in contrast, restores *any* status quo ante regardless of whether or not it displayed any such characteristics.

Consider, first, end-state theories. They assert that we should provide compensation because the pattern embodied in the antecedent distribution was the right one. For it to have been upset was wrong and for it to be restored would be right. Compensation would do that. Hence we should compensate.

Notice, however, that in practice the payment of compensation is independent of any judgment about the justice of the antecedent pattern of distribution. People claiming compensation need not prove, to win their case, that damage done to them exacerbated relative deprivation or social injustice. They need prove only that they were harmed.

The law of torts protects rich and poor alike. People may claim compensatory damages whenever they can prove that they were harmed by another's tortious conduct. To win their case, they do not have to prove they need the money. Those making ostentatious displays of donating their damage awards to charity were no less entitled to receive them than were those who have to use the money to buy basic necessities.⁷ Those accused of torts cannot say, in their defense, that they took only from those richer than themselves.⁸ Current tax laws may imply that an ideally just pattern of distribution entailed a transfer of 10 percent of the income of those in high tax brackets to those in low tax brackets; but the moral superiority of that pattern

⁶ By that, I mean the practice of providing compensation – understood as the full and perfect equivalent of what was lost – to right any and all wrongful damage to persons and property. Although that ideal is rarely realized perfectly, it is the ideal practice rule rather than the inevitably imperfect implementation of it that a theory of compensatory justice must strive to rationalize. Compensation is often intertwined with various other practices, such as deterrence of harm-causing activities, via liability rules in torts. Many of the peculiarities of notionally compensatory practices, inexplicable on my account, might be explained by reference to those other, competing considerations. Still, I trust that there are enough “clean” cases of compensation to build a theory around those paradigmatic instances.

⁷ Although there have been pleas, so far largely unmet, for tort damage awards to be more explicitly redistributive (Abel 1982).

⁸ Likewise, in contract law, we may for distributional reasons refuse to countenance certain sorts of contract (entailing usurious interest rates or slave wages, e.g. [Kronman 1980]); but such distributional considerations primarily constrain what sorts of contracts will be considered valid in the first place, rather than excusing breach of an otherwise valid contract. Poor debtors are not excused, by reason of their relative poverty alone, from repaying debts owed to the rich.

does not excuse the poor in stealing from the rich, provided they take no more than the 10 percent ideally coming to them.

In the law of eminent domain, similarly, the government is obliged to compensate property owners whenever it takes their property. That obligation is independent of any judgment about the justice of their holding that much property in the first place. Inheritance tax laws may imply that it is unjust for property to be concentrated in so few hands.¹¹ We would nonetheless be required to compensate a landowner for property seized during his life, even if we had intended to take it without compensation upon his death.

Or, again, social security benefits are (as has been seen in Chapter 12) characteristically earnings related. Workmen's compensation, unemployment compensation, accidental injuries compensation, criminal injuries compensation and the like all aim to compensate people, *inter alia*, for whatever earnings they have actually lost. In consequence, those programs all pay people more the more they earned. That is independent of, and in certain ways flatly contrary to, public judgments made in other contexts about the justice of income differentials.¹² The results may be paradoxical, but the practice of paying compensation in this fashion is well nigh universal.

None of this, of course, is to say that we refuse ever to countenance inquiries about the justice of patterns of holdings. Those inquiries must merely proceed along a separate track. The duty to pay compensation for damage done is independent of whether people have relatively too much or relatively too little by way of present endowments. That, in turn, implies that the reason we insist on compensation does not have crucially to do with the correctness of the pattern of the distribution that the compensation would restore.

Consider, next, historical entitlement theories. Those assert that we should compensate because people have been deprived of what is, by historical entitlement, rightfully theirs. To upset that distribution was wrong, and it should be restored. Compensation would do that. Hence we should compensate.

But that is not the way compensation works in practice. Consider

¹¹ Process-based, historical entitlement might allow a more nuanced interpretation of inheritance tax policies (e.g., that it is not wrong for the rich to enjoy their riches, it is merely wrong for them to pass them on to others). In the end-state terms here in view, though, the pattern of holdings is all that can matter. There is no place for considerations about how people came by their riches – through their own efforts or through bequests – to enter these calculations.

¹² We know from its tax policies that the government itself thinks that a more equal pattern of income distribution would be preferable and that it is the government's job to promote it. Yet the very same government, through its compensation policies, sets systematically about reproducing the same nonideal pattern of income distribution which it tries to correct through its tax policies.

the paradigm of tort law. There is no requirement that courts mount a title search to ensure that people have historically clear claims to their property before requiring tortfeasors to compensate them for damage to it.¹¹ As the Reporters for the American Law Institute's *Second Restatement of the Law of Torts* comment, "The important thing in the law of torts is the possession [of property], and not whether it is or is not rightful as between the possessor and some third party."¹⁴

The tortfeasor's duty to compensate is not defeasible, even upon production of clear evidence of force or fraud in the history of the plaintiff's acquisition of the damaged property.¹⁵ J. D. Rockefeller may have engaged in some dubious practices, legally as well as morally, in building the Standard Oil empire. That fact does not make it any less of a tort for me to spraypaint graffiti on that company's gas pumps, nor does it do anything to reduce the damages I have to pay in compensation for cleaning them up.

Similarly, under the law of eminent domain just compensation is owed to those who possess (and hence have an interest in) property, whatever the strength of their legal title.¹⁶ Or, again, under various social insurance schemes such as workmen's compensation my entitlement is contingent merely upon the requisite contributions having been made. How I or my employer came by the money used to fund those contributions is irrelevant to my entitlement. Or, for yet another example, my entitlement to unemployment compensation is contingent merely upon having lost my job. Whether that job was rightfully mine or whether I bribed someone to get it is irrelevant.¹⁷

None of this is to say that we refuse to look into the historical basis of people's titles to their property. It is simply to say that those inquiries, too, must proceed along a separate track. The duty to compensate for damage done is independent of those other inquiries. That, in turn, suggests that the reason we insist upon compensation does

¹¹ It is not as if questions of title never matter. Thieves brash enough to sue for damage to automobiles they have stolen are unlikely to persuade many juries to decide in their favor. My point is just that questions of title do not always enter.

¹⁴ Prosser 1965, sec. 328E, comment. Those "who can recover for private nuisance" include first and foremost "possessors of the land," a term which the Reporters pointedly add "applies to adverse possessors [i.e., those who have no title to it] as well as to those rightfully in possession" (Prosser 1965, sec. 821E, comment c).

¹⁵ The *Restatement (Second) of the Law of Torts*, again, specifies: "One is not barred from recovery . . . merely because at the time of the interference he was committing a tort or a crime or, in the case of an interference with his title to or possession of lands or chattels, because it was tortious or illegal for him to have the title or possession" (Prosser 1965, sec. 889).

¹⁶ American Law Institute 1986, sec. 609.

¹⁷ Unless, perhaps, the reason I lost the job was in penalty for having obtained the job improperly in the first place.

not have crucially to do with the historical rightness of the distribution that compensation would serve to restore.

The postulate that compensation is justified on the grounds that it restores a substantively right distribution cannot therefore be sustained. Different theories exist as to what makes distributions substantively right. But all would sensitize us to particular right-making characteristics that defensible distributions must have, whereas compensation is in practice not systematically sensitive (to the point of being almost systematically insensitive) to such considerations.

B. Righting wrongs

The preceding arguments justified compensation in terms of claims about the substantive rightness of the status quo ante that it would restore. A second set of arguments justifies compensation in terms of the wrongness of the process by which that antecedent distribution has been upset.

This analysis has a certain appeal. Without necessarily joining historical entitlement theorists in claiming that outcomes are substantively right if produced through certain processes, we can nonetheless agree that outcomes are wrong if produced through processes involving force or fraud, for example. Righting procedural wrongs may be justifiable independently of any theory of the substantive rightness of the outcomes thereby produced.

This analysis comes naturally to those approaching compensation by way of torts. Those, by definition, are wrongs; and their wrongness lies not in the substantive wrongness of the outcomes produced but rather in the wrongness of producing outcomes in those ways. Compensatory damages right wrongs, in torts, not in the sense of restoring substantively right distributions but rather in the sense of canceling the effects of wrongful styles of intervention in others' affairs.¹⁸

True though that analysis of "righting wrongs" may be of tort damages, it is simply untrue of various other modes of compensation equally important in today's world. Consider especially no-fault compensation systems. Both workmen's compensation and New Zealand's more general accidental injuries compensation are paid to victims, independently of any inquiry into fault or blame or wrong, whether on the part of the victims or of anyone else. Those modes of compensation aim not so much at righting wrongs as eradicating evils. Compensation, then, is a mixed bag. In some cases (notably, tort damages

¹⁸ Similarly, just compensation for takings of private property under powers of eminent domain has historically been analyzed as part and parcel of "due process" (Corwin 1911, p. 378; Grant 1931).

and criminal injuries compensation), it is clearly designed to right procedural wrongs. In others (notably, workmen's compensation and accidental injury compensation), clearly it is not. And still other cases fall between these two poles.¹⁹

The only way to save the analysis of compensation as justifiable righting of procedural wrongs, at this point, is to argue that tort-style compensation is somehow paradigmatic. An argument along these lines might be mounted.²⁰ But there is always going to be something deeply unsatisfying about such an analysis. Holding some cases of compensation to be paradigmatic and others deviant amounts to little more than attempting to explain away our inability to explain half the cases the theory is supposed to cover. Instead of giving a rationale for compensation, that amounts to rationalizing our failure to provide a rationale for a great deal of compensatory practice. Were that the best we could do, it might be good enough. As I hope to show in Section IC, an alternative rationale for compensation is both more complete and more convincing.

C. Underwriting reasonable expectations

Having forsaken rationales for compensation based on the substantive rightness of the previous situation or on the wrongness of the way in which it has been upset, we are left looking for its justification in some more formal feature of the status quo ante. The substantive content, historical background and distributional pattern of the status quo ante are all apparently irrelevant to our decision to provide compensation

¹⁹ Unemployment compensation is nearer the latter pole. Claiming unemployment benefit is importantly different from mounting a claim for wrongful dismissal. But there is a tinge of fault-based logic in the fact that a claim for unemployment compensation might, in many places, be denied if the claimant were himself responsible for his own unemployment (by, e.g., having voluntarily resigned his post or having given an employer good grounds for dismissal). While unemployment compensation is not compensation to right a wrong, therefore, claims for it might be defeasible on the basis of a wrong of a certain sort on the claimant's part.

²⁰ Historically, workmen's compensation became no-fault less for reasons of high moral principle than for pragmatic reasons. Establishing fault often proved prohibitively expensive, thereby blocking otherwise meritorious claims. Contemporary concerns for extending the no-fault principle are often similarly motivated pragmatic concerns. These no-fault schemes of compensation might therefore really be designed to right *presumptive* wrongs which would cost more than it is worth to prove conclusively (Freeden 1978; Atiyah and Cane 1987, chap. 21). But that still makes little sense of trends to extend compensation, nowadays, to cases of (blameless) accidents and, increasingly, of disease involving virtually no human agency at all (Atiyah and Cane 1987, chap. 20; Stapleton 1986).

to restore it. What we apparently need is an argument to the effect that we should restore it, merely because the previous state *was* the status quo ante.

No simple argument along those lines is going to work. The mere fact that "x existed" is no reason for supposing that it should have existed. It provides still less of a reason for supposing that it should be re-created once it has *ceased* to exist.²¹ The sheer fact that the status quo ante was the status quo ante cannot justify a policy of restoring it.

There might, however, be a way to derive the requisite justification from a more complex set of facts, all connected somehow to the fact that the state being restored was indeed the status quo ante. The further propositions that we need to complete this more complex version of the nonsubstantive, nonprocedural rationale for compensation, are three:

1. People reasonably rely upon a settled state of affairs persisting (or, anyway, not being interrupted in the ways against which compensation protects them) when framing their life plans.
2. That people should be able to plan their lives is morally desirable.
3. Compensation, if sufficiently swift, full and certain, would restore the conditions that people were relying on when framing their plans, and so allow them to carry on with their plans with minimal interruption.²²

There is no need to belabor the central ethical premise of this analysis (proposition 2). The moral desirability of people being able to frame and follow through on their life plans can be defended in a variety of familiar ways. One points to the value of autonomy. Another points to unity and coherence as a source of value in people's lives. Yet another points, in indirect, rule-utilitarian fashion, to the value that we all derive from being able to anticipate what each other is going to do.²³

All I here need add to those standard accounts is the further proposition that, however important it may be that people should be able

²¹ One problem is that multiple candidates appear for the role of status quo ante, each corresponding to a different past period. We have no reason for favoring any one over the others. Another problem is that arguments for protecting the status quo ante often illicitly turn on arguments for protecting the status quo as such. Yet once the status quo ante has been upset, there will have been established a new status quo, itself equally deserving of protection under that implicit principle.

²² This expectation-protecting rationale appears often in discussions of compensation; see Feldstein 1976a, p. 124; Cordes and Goldfarb 1983, p. 356; and Michelman 1967, pp. 1211-13 and *passim*.

²³ These are discussed at length in Chapters 11 and 12.

to frame life plans for themselves, it is all the *more* important that they be able to follow through with those plans once they have embarked upon them. The point of planning is not the act of planning per se, but seeing those plans come to fruition. Whether you tell an autonomy-based story in terms of sunk psychological costs or an indirect-utilitarian story in terms of reliance upon established patterns in others' behavior, it proves to be especially important that people be able to follow through with their plans once they have set off down some particular path.

Not much elaboration is required on the third proposition, either. The function of compensation is straightforwardly to restore the status quo ante. It will serve its expectation-preserving purposes, however, only if certain further conditions are satisfied. The compensation must be *complete*, giving people back the full and perfect equivalents of what they lost, if they are to be able to carry on as before. It must be *swift*, restoring them promptly to avoid damaging interruptions to their ongoing projects.²⁴ And it must be *certain*, allowing no doubt that compensation will be forthcoming, so they can plan with confidence. Should any of these conditions fail to be satisfied, people's plans would be at risk of irremediable harm.

The first proposition will presumably be the most contentious. All kinds of objections are raised against theories tracing entitlements to "reliance" or "expectations" more generally. Hume's theory of property rights and the reliance theory of contractual obligation are both criticized on the ground that people can form completely baseless expectations and rely on utterly unreasonable assumptions. That someone expects or has relied on something is not conclusive, morally, in the absence of some reason for supposing that good grounds exist for the reasonableness of that expectation or reliance.

One response to such an objection is this. Nothing is unreasonable,

²⁴ This analysis is perfectly consistent with, and borne out by, the practice of compensation. On the one hand, we think it important that compensation should be paid promptly, and we regard it a scandal that the average time between injury and tort judgment is three years (Atiyah and Cane 1987, p. 272). On the other hand, we also seem to think that after a certain period of time, no compensation need be paid at all. On tort claims the statute of limitations typically specifies a few years; on social security claims, such as for workmen's compensation, the time limit for claiming is usually a few weeks or months. The most standard reason given for that practice has to do with problems of amassing reliable evidence long after the event, which may be a powerful reason in some cases (e.g., torts) but is weak in others (e.g., workmen's compensation, where the factory safety officer's log or hospital's records are utterly reliable long after the event). The most satisfactory reason for the practice, as I suggest in Chapter 12, seems to be that compensation is supposed to avoid interruptions to people's life plans; and such compensation would have no point long after the event, because by then that interruption would already have occurred.

statistically, in predicting that a settled state of affairs will persist. Induction is not *that* unwarranted; the future really is only marginally different from the present, very much more often than not. But suppose the status quo embodies a pattern of systematic and entrenched injustice. Statistically we would have good grounds for predicting it to persist, but we would have no grounds morally for expecting (still less, for demanding) that it do so. We must therefore say something further about *what sorts* of interruptions to people's plans we propose to remedy through compensation.

We do not try to reverse all changes in people's plight. We compensate only in certain, well-defined situations: accident, crime, disability, unemployment, tortious wrong and so forth. That catalog seems to suggest that we compensate to restore the status quo ante only if it has been upset either (1) in ways people had no reason to expect or (2) in ways people had a reason to expect not to occur.²⁵

As usual, that formulation intentionally straddles the statistically unpredictable and the morally unacceptable.²⁶ That allows us to deem "reasonable" – and hence, by my formula, compensatable – expectations which are well founded morally but not statistically. It would be perfectly reasonable, morally, for people to expect crime not to happen, even if they knew, statistically, that it often did.²⁷

It also follows from my analysis that certain sorts of alterations to the status quo should not be compensatable. It is unreasonable in a moral (if not, alas, in any statistical) sense for thieves to expect to retain their booty or for monopolists to expect to continue enjoying the fruits of their privileged market position. No compensation is due them to make up for any interruption to their life plans when those things are taken from them by due process of law.²⁸ The same may be true of certain risky ventures and unsettled situations. As a matter of policy, we guarantee the security of bank deposits but we do not underwrite the value of stock portfolios. The reason – consistent with the reason I offer for compensation generally – is that people have no good grounds, of a statistical or still less a moral sort, for expecting

²⁵ Clause (2) goes some substantial way toward subsuming the model of righting wrongs analyzed in Section IB. That is acceptable, since the arguments offered against that model merely served to suggest that it is at least a partial account, true of some but not all cases of compensation. If clause (2) here is taken as subsuming that model, clause (1) can be taken as providing the rest of the story needed for a complete account.

²⁶ Hart and Honoré 1985, chaps. 2 and 3. Mackie 1955. Nozick 1972, p. 112.

²⁷ Though at the margins statistical expectations trump moral ones. Perhaps walking alone in Central Park after dark, or not locking your doors in Detroit, really is "bringing the harm on yourself" in a way that makes it noncompensatable, however immoral the crime itself might have been.

²⁸ Michelman 1967, pp. 1235–9. Cordes and Goldfarb 1983, pp. 364–5.

high business profits over the long run, in a way they did (even before the advent of the Federal Deposit Insurance Corporation) have good statistical grounds for expecting banks usually not to fail.²⁹

It is sadly true that we do not always pay everyone all the compensation that on my account they truly deserve. Sometimes compensation would necessarily be so grossly inadequate that compensation would constitute a travesty. Where harms are truly irreparable, and the loss truly irreplaceable, little more than token compensation is offered.³⁰ Other times the problem lies, purely pragmatically, in figuring out who should pay the compensation.³¹ Thus it seems at least arguable that the ethical point of compensation always lay in stabilizing expectations, and only purely pragmatic considerations ever stopped us from going the whole way toward that ideal.

D. Interim conclusion

The true justification for compensation lies not in the substantive content of the status quo ante that it restores, or in the wrongfulness of the process by which that status quo ante was upset, or in the fact that the situation being restored had indeed been the status quo. The justification lies instead in a complex set of facts about the way in which people had been reasonably relying upon the settled status quo

²⁹ The standard *reductio ad absurdum* of the reliance theory of contractual obligations goes like this: I could write H. Ross Perot, saying that I was relying upon the expectation that he would send me \$10,000 to prevent foreclosure on my mortgage; and so long as the purported reliance is genuine Perot would have an obligation to send the money, even though he had made no prior commitment to do so and even though my reliance upon his doing so was baseless (Fried 1981, p. 10). Whether or not this is a good *reductio* of the reliance theory of contract law (cf. Goodin 1985c, pp. 42–52), it is no counterexample to the present expectation-based analysis of compensation. Expectations in my theory must be tied to some status quo ante, in a way that that expectation about Perot's benefaction clearly was not.

³⁰ D. Harris et al. 1984, p. 90. See further Chapter 11, Section 2, in this volume.

³¹ Consider the history of compensation schemes. Tort damages came first, because it was clear who should pay; workmen's compensation and the like came next, once we had worked out principles of insurance allowing us to charge everyone involved in risky enterprises premiums proportional to their contributions to the risk of accident. The next steps are presumably to extend the same insurance concepts to embrace compensation for disease, economic dislocation and so forth. The ease with which compensation schemes worldwide tended to evolve in these directions suggests that the reason for beginning where we did, and pausing where we have, was only ever pragmatic. It is not as if we ever thought that, on principle, to deserve compensation you had to have been injured in some particular way. We merely found it easier saying from whom people injured in those ways should recover.

ante persisting in much the same shape into the future when framing their life plans.

This analysis presents no more than a *prima facie* justification for compensation. It provides one reason for compensation restoring the status quo ante, but that reason is not necessarily conclusive. Countervailing reasons, which may well prove stronger from time to time, may argue for altering the status quo in certain respects – even at the cost of upsetting expectations. The argument I give for compensation does not say that that must never be done, merely that it will always carry some moral cost.

My case for compensation also implies, however, that that cost is variable rather than constant. The status quo can often be altered at little cost, and sometimes at no cost at all, in terms of upsetting people's expectations. Insofar as one of these less-cost strategies is pursued, we have less reason (in the limiting case, no reason at all) in favor of compensation or against redistribution.

II. THE COMPATIBILITY OF REDISTRIBUTION AND COMPENSATION

In this section I am concerned less to argue for redistribution than to block certain sorts of arguments against it. As I said at the outset of this chapter, compensatory justice seems systematically to trump principles of redistributive justice. How powerful that trump is, though, depends on how principles of compensatory justice are justified. If the reason for restoring the status quo ante had to do with its substantive rightness either because of its pattern or of its history, then any redistribution (which necessarily deviates from that right distribution, also) would be condemned. But redistributivists need not fear. Compensation, as I have shown, cannot find justification in any such facts about the substantive content or historical pedigree of the status quo ante.

Redistribution necessarily alters the status quo. If the reason we should compensate were that it was wrong to alter the status quo, because the distribution being altered was substantively just by some standard or another, then redistribution would be wrong for precisely the same reason that compensation would be right. But if the reason we should compensate is not that it is wrong for the status quo to have been altered, but merely wrong for it to have been altered in certain sorts of unanticipatable ways, then the conflict between redistribution and compensation is erased.

On the expectations-based argument, what is wrong is not altering the status quo but rather altering it unpredictably. Therefore redistribution that alters the status quo would be perfectly permissible, just

so long as it was done in a predictable manner. This is just to say that redistributions ought to abide by something akin to rules of natural justice. In economic policy just as in criminal law, public affairs ought so far as possible to be conducted according to known rules.¹³ So long as people know well in advance when and how redistributive policies will affect them, redistributions cannot come as a "bolt out of the blue" wreaking unreasonable (because unanticipatable) havoc on their lives.

Standing rules of distribution of precisely that sort are written into our tax codes. The progressive income tax has long been on the books. Its rates change marginally from time to time, but everyone has had more than adequate warning of the government's general intention to tax higher incomes more heavily than lower; indeed, given how little the real rate of tax paid varies from year to year, people have even had adequate warning of the rough magnitudes involved. Since those laws predate any of us entering the work force, we all ought reasonably be expected to have framed our life plans taking those redistributive measures into account. We would have no grounds for complaining about an unanticipatable interruption in our plans when those tax bills fall due, had we failed to do so.

My argument on that score presupposes, first, that people can reasonably be expected to take the previously announced policies of their government into account when forming their plans. That seems minimally contentious, at least insofar as those policies constitute "settled intentions" of the government.¹⁴ My case further presupposes that governments can make and stick to redistributive policies that stretch well beyond any individual's planning horizon. Governments surely do so; they engage in long-term economic planning, investment in basic research and development and so forth on a far longer time horizon than that of the typical individual or firm.¹⁴

¹³ Fuller 1964.

¹⁴ And, we might want to add, insofar as those policies are not actually immoral. Whether that proviso is strictly necessary is something of an open question. It seems unreasonable for people to proceed with their life plans in the expectation that the government's settled redistributivist intention, however immoral, will not be acted on, in the same way as it is unreasonable for people to persist in strolling through Central Park at night long after it has become a den of thieves. The mugged stroller, like the unjustly treated citizen, can hardly profess *surprise* (as distinct from moral outrage) when pulled aside by the thief or the police officer. But the proviso hardly matters, since advocates of redistribution will in practice have to have something persuasive to say in its moral defense.

¹⁴ In democracies, periodic elections ensure that governments have short life expectancies. But one government's redistributive plans are substantially retained, much more often than not, by its successors, even in cases of dramatic regime changes (Rose and Peters 1979, pp. 115, 263).

The main thrust of my compatibilist claim thus has to do with an *ongoing* system of redistribution. Where redistribution is conducted according to settled principles of long standing, we have no reason to fear that it will do any real violence to the sorts of moral values that compensation is designed to serve. That of course leaves the problem of transitions. How are we to institute those rules of redistribution in the first place?

Problems of transitions are genuinely difficult in practical terms and genuinely fascinating intellectually. They have, accordingly, generated a large literature and much discussion.¹⁴ While I have no compelling solution to the problem of transitions, neither is that as much of an omission as it might seem. Transitions, after all, are merely transitional – short-lived, episodic and hence relatively inconsequential – compared to the more protracted regimes on either side of them. I am, therefore, much more concerned to give a proper account of those settled states than I am to give a good account of transitions.

Still, my analysis does offer some potential solutions to the problem of how to shift from less to more redistributive regimes with minimal interruption to people's life plans. One solution is simply advance notice.¹⁵ Delayed implementation of this sort might help to cushion redistribution's blow to your life plans. The problem from the redistributivist's perspective is that it does so by reducing the redistributive effect of the policy, at least in the short term.

Another alternative would be to redistribute without warning, but then to compensate losers straightaway in some other currency. Economists, anthropologists and political theorists alike observe that we tend to be "specific egalitarians," worrying more about the distribution of some commodities than of others or of money in general.¹⁶ Insofar as we are, we can take advantage of this curious fact to redistribute what matters and to compensate with what does not. In this way, redistribution might once again be squared with the deeper values that compensation is supposed to serve.

Bringing about redistributions in the ways I propose, either in the transitional case or in an ongoing system of redistribution, carries costs of various other sorts. I would not want to deny that. I merely say that if you are worried about the sorts of moral consideration that motivate compensation – things like protecting reasonable expecta-

¹⁴ See Hochman 1974, Kaplow 1986 and sources cited therein.

¹⁵ Tullock (1978, p. 53) commends the example of nineteenth-century reforms abolishing various sinecures in the British civil service only upon the death of their incumbents; others offer more modest suggestions along similar lines. See similarly Feldstein 1976b, pp. 98–9; cf. Goldfarb 1980, pp. 29–30.

¹⁶ Tobin 1970. Douglas and Isherwood 1979. Walzer 1983.

tions and life plans built around them – then there are ways to arrange your redistributive scheme to protect those values.

That may or may not be the optimal mode of redistribution, all things considered. Consider, for example, the familiar argument of economists that if we want to redistribute we should do it through unanticipated, once-and-for-all, lump-sum transfers in order to minimize the disincentive effects and consequent efficiency losses.¹⁸ The redistributive “bolt from the blue” that Pigou and his followers recommend on efficiency grounds would wreak havoc with people’s lives. Indeed, it is crucial to the success of his efficiently protecting scheme that the redistribution be utterly unpredictable. Whether the efficiency gain would be worth the cost in terms of disruption to people’s lives is an open question.¹⁹

Happily, though, it is a question that I do not have to answer here. My compatibilist thesis does not hold that redistribution of the form I suggest is compatible with *every* value that we may wish public policy to promote. My claim is merely that, perhaps surprisingly, such redistribution is thoroughly consistent with the purposes that compensation should serve. So long as that is true, arguing for compensation is not necessarily arguing against redistribution.

III. THE MUTUAL ENTAILMENT OF COMPENSATION AND REDISTRIBUTION

The point of my previous argument is that redistribution is consistent with compensation. That is to say, it does not necessarily offend against compensation’s underlying moral principle, properly understood. Perhaps we can go further and say that redistribution actually implies or is implied by compensation in certain circumstances. Maybe the same moral principle that demands compensation sometimes also demands redistribution, or vice versa.

I offer these speculations more tentatively. The central argument of the chapter is the compatibilist thesis already established; and that argument is perfectly capable of standing on its own, without any support from the ones I am about to set out. Still, if these further extensions also go through, I will have succeeded not only in blocking

¹⁸ Pigou 1932, pt. 4, chap. 9.

¹⁹ For some estimate of the magnitude of the effects, consider the calculations of Danziger et al. (1981, p. 1019). Contemporary U.S. transfer payments, practiced in a markedly anti-Pigovian way, seem to produce a 4.8 percent reduction in labor supply, owing to disincentive effects, in exchange for a reduction of 75 percent in poverty and a 19 percent reduction in the Gini coefficient of income inequality.

an apparently powerful argument against redistribution but in actually advancing the positive case for it.

A. *Redistribution implies compensation*

The thought with which this chapter began is that compensation and redistribution are implacable foes. It has now been shown that the two can actually be reconciled. But showing that the two are not necessarily enemies is still far short of showing, as I shall now attempt, that they are necessarily friends.

The first leg of that argument proceeds by a *reductio*. Suppose we harbor egalitarian ideals. Suppose, further, that some rich person suffers accidental or even tortious damage, of the sort for which we think people ordinarily should be compensated. If we are egalitarians, why bother? Under our redistributive program, rich people are scheduled to have a certain amount taken away from them anyway. Dame Fortune or tortfeasors have simply spared us the trouble.⁴⁰

That way of thinking is plainly crazy. No one supposes that insurance companies should refuse to pay out on fire insurance policies just because the house that burned down was a rich person's; no one even thinks that, as a matter of public policy, houses should be uninsurable above a certain value.⁴¹ The same seems true of all the other hazards that face us in contemporary life, and of all the other mechanisms that we have for compensating people when they strike.

What is fundamentally wrong with such propositions is just this. We do not want to redistribute by accident. Redistribution is a matter of policy, not of happenstance. Those who hold redistributivist ideals invariably demand *intentional* redistribution. They want to produce a certain pattern of holdings, or they want to rectify certain historical wrongs. But they wish to do so systematically, not randomly.

In the same way that redistributivists think it unjust for some people to get rich by sheer luck, so too must they agree that it would be unjust for some but not others to be relieved of their undeserved riches by the sheer bad luck of their being the uncompensated victims of accidents or injuries. No one is in favor of "capricious redistribu-

⁴⁰ Perhaps the most explicit expression of this thought comes from the "new welfare economists," who left compensation merely hypothetical precisely on the grounds that insisting upon the payment of actual compensation would unjustifiably lock in the existing distribution and arbitrarily preclude the possibility of redistribution (Kaldor 1939, pp. 550-1; 1946-7, p. 49; Hicks 1939, pp. 711-12; Scitovsky 1941).

⁴¹ Furthermore, that does not just reflect some general prejudice against ever preventing people from insuring against whatever they want. Some things, such as punitive damages in torts, are already uninsurable.

tion."⁴⁴ The upshot is that redistribution-in-the-large seems to imply compensation-in-the-small.

The compensation thereby endorsed by redistributivist ideals is highly qualified. Money given to the rich person whose house burned down may well be subsequently taxed away in the course of a larger redistributivist program. And the losses intentionally inflicted upon the rich by a scheme of redistribution can themselves hardly be compensatable. So the sort of compensation implied by redistribution is both provisional and partial. What is surprising is not that it is shaky in these ways but, rather, that the implication is there at all.

B. Compensation implies redistribution

In similar fashion, the goals that compensation is supposed to serve might actually imply a certain measure of redistribution. Once again, the implication will certainly be partial and may well be weak. The surprise is simply that it should be there at all.

In Section I, it was shown that compensation standardly strives to restore some status quo ante. That general rule has one conspicuous exception, though. We run programs of "disability compensation" of various sorts – invalid pensions, attendance allowances or invalid care allowances, mobility allowances and the like.⁴⁵ Furthermore, we offer those benefits even to the congenitally handicapped. In providing mobility assistance to the congenitally handicapped, however, we are not restoring them to some status quo ante in which they were able to walk. Their handicaps being congenital, their impaired mobility has been lifelong. That makes this case very unlike the ordinary practice of compensation.

What we are doing here is not restoring the congenitally handicapped to some status quo ante. Rather we are bringing them up to a standard that, while normal for the species, is one that those particular individuals never actually enjoyed.⁴⁶ Worthy though that practice

⁴⁴ Michelman 1967, pp. 1217–18. Similarly we insist upon just compensation for property taken under the power of eminent domain, not so much for Epstein's (1985) contract/tort principles (if you want to use another's property, you must pay for the privilege) as for public-finance ones that essentially "bar the Government from forcing some people alone to bear public burdens which, in all fairness and justice, should be borne by the public as a whole" (Black 1960, p. 49). Just as it would be inequitable to impose differential taxes on people in the same income class, so too in the takings case would it be similarly inequitable to impose differential sacrifices on people merely because they happen to own something that the public requires.

⁴⁵ tenBroek 1966. D. Harris et al. 1984, pp. 4–12. Haveman et al. 1984. Atiyah and Cane 1987, chap. 16.

⁴⁶ Culyer 1974, pp. 22–3.

may be, it cannot be justified in the same way as ordinary compensation. Bringing up to some minimal norm of human existence people who never previously enjoyed it does certainly not "restore" anything to them. In terms of the rationale for compensation developed here and in Chapter 11, their life plans have not been interrupted by any dramatic changes.

These people might, however, suffer a closely related problem. Without that "minimal baseline norm" secured, these people may be unable to frame and to follow life plans at all. If it is morally desirable that people should do so – as one crucial step in my argument for compensation asserts it to be – then perhaps redistributive transfers to underwrite that "minimal baseline norm of human existence" are themselves morally mandated by the selfsame principle. It may be confusing to call it "compensation." But the confusion is minor, since redistribution of that sort serves the same moral goal as compensation itself.

Exactly where the "minimal baseline norms" should be set for what goods should probably be left open. In general, what this "preconditions of agency" style of argument implies is that those norms should guarantee everyone enough of what they need to be able to plan their lives at all. What that amounts to is an empirical matter. It seems empirically likely, however, that the gravest forms of deprivation remedied by the redistributive transfers of the welfare state – gross poverty, protracted unemployment, persistent illness and homelessness – would all qualify as major interferences with any planning at all. They therefore qualify for relief under my principle.⁴¹

The "preconditions of agency" argument is most persuasive when the resources involved are, literally, preconditions. That is to say, the move is most persuasive when it is literally impossible to frame and follow any life plans at all unless those resource needs have been met. But that is a limiting case that will rarely occur. The handicaps for which programs to aid the disabled compensate (impaired mobility and the like) reduce one's options. But they hardly preclude one from conceptualizing plans, nor do they usually even reduce one's options down to a single possibility. The same is arguably true of poverty and most of its corollaries.⁴²

If "preconditions" is too strong, "preoccupations" may serve almost as well. The reason we should redistribute resources to meet

⁴¹ Plant 1988. Waldron 1993a, pp. 225–49, 309–38.

⁴² It may well be that food, clothing and shelter – and the basic income required, in market societies like our own, to secure them – are "primary goods," necessary for whatever else one wants to do. Still, one can always cast one's plans in contingent terms, stipulating what one would do were one's basic needs satisfied. Such primary goods may be preconditions of efficacious acting, but they are not preconditions of planning *per se*.

people's basic needs is that without those resources the psychological prerequisites of planning one's life are lacking. Almost inevitably, people who do not know where their next meal will come from or where they will sleep tonight will find those concerns completely absorbing. Preoccupied with how to satisfy their immediate needs, they are incapable of thinking much beyond that. If we regard it as desirable for people to frame and follow plans of a larger sort for their lives, we ought therefore do what we can to remove those barriers to such longer-term planning.

This argument claims that redistribution may be needed to enable some people to plan in the first place, whereas the previous arguments maintained that we need to compensate people to allow them to carry on with reasonable life plans once embarked upon them. The difference between facilitating the framing of plans and facilitating follow-through on projects in progress may well matter enormously, morally. Logically, perhaps, making plans comes before carrying them out. But the disruption of projects-in-progress matters far more, both phenomenologically and consequently morally, at least for any morality that takes people's self-respect seriously. That blunts the force of preconditions-of-planning style arguments for redistribution, which is one reason I prefer to emphasize the arguments of previous sections and to put these propositions more tentatively.

Another reason is that admitting a connection between autonomy and resources is a double-edged sword. That connection having been made, when resources are redistributed away from the rich, they can then complain of a loss of autonomy to them that is strictly analogous to the gains in the autonomy of the poor that they experience from the resources being redistributed toward them.⁴⁷

I do not think that that complaint is compelling. If resources yield diminishing marginal autonomy just as they do diminishing marginal utility, redistribution will do more to promote the autonomy of the poor than to decrease the autonomy of the rich. Moreover, there is a world of difference, in autonomy terms, between not being able to plan at all and not being able to carry out fancy plans for some highly luxurious existence. But those are larger arguments than can be pursued here. And the arguments of the first two sections are, for the purposes of this chapter, more than enough anyway.

IV. CONCLUSION

My analysis explains the apparent paradox of running compensatory side-by-side with redistributive policies. Paying some relatively rich

⁴⁷ Jones 1982.

victim compensation that redistributive measures will claw back may seem to amount to giving with one hand and taking back with the other. But there is method in that seeming madness.

If we think it is morally desirable to ensure that people are able to plan and organize their lives in a sensible fashion, we must be systematic about both compensation and redistribution. Just as people must be able to count on compensation whenever harmed in certain ways, so too must they be able to count on redistributive policies working in a similarly relentless fashion, no matter whether they are on the giving or the receiving end of those policies. Erratic compensation would be profoundly unsatisfactory. For the same reason, we should be profoundly unsatisfied with redistribution erratically taking only from those unlucky enough to suffer harms and rewarding only those lucky enough to benefit from their misfortunes. Both compensatory and redistributive policies can in this way be seen as manifestations of broadly the same principle, a principle of subjecting the contingencies that buffet individuals' lives to some sort of rational public control.⁴⁸

⁴⁸ Earlier versions of these were discussed in London, Norwich and New Orleans. I am particularly grateful for the comments, then and later, of Elizabeth Anderson, Brian Barry, Debbie Fitzmaurice, Martin Hollis, Sheldon Leader, Saul Levmore, Onora O'Neill, Morris Perlman and Albert Weale.

Chapter 14

Basic income

Proposals to pay everyone in society a "basic income" are, on their face, straightforward. The same sum would be paid, absolutely unconditionally, to everyone in society – regardless of their work performance, income, needs or any other personal characteristics.¹ Of course, people might supplement this basic income in various ways, typically with earned income. But it would make no difference to the "basic income" you receive whether you were young or old, in work or out of it, willing to work or unwilling or unable to do so, married or single, had dependents or not, or whatever. Every individual in society would be guaranteed the same basic sum, regardless.

Many arguments can be offered in support of such schemes of income support. Egalitarian aims are served by paying everyone the same sum, libertarian ones by giving them cash rather than end-use goods, civil-libertarian ones by making the grant unconditional and hence immune to official inquiries concerning people's personal circumstances.² True and important though all that may be, I want to focus in this chapter on another equally compelling but less familiar rationale.

The crux of my case is that schemes paying everyone an unconditional basic income are *less presumptuous* than more conditional programs of income support. Not only are they less prying and intrusive, less demeaning and debasing. More to the point of my present argument, they also simply make fewer assumptions and presumptions about those whom they are aiding. That in turn makes schemes of basic income more *efficient*, in one important sense, than more conditional schemes of income support.

¹ Van Parijs 1986; 1991; 1992, chap. 1; and van der Veen and Van Parijs 1987.

² Baker 1992. Steiner 1992; Friedman and Friedman 1980, chap. 4. Reich 1963; Handler 1979; Goodin 1988, chap. 7. See further Van Parijs 1992, esp. chap. 1, and van der Veen and Van Parijs 1987.

Critics of schemes of generous social benefits often complain of their alleged "inefficiency," which in those contexts amounts to saying that we spend too much on social welfare goals and pay too high a price in terms of other goals in pursuit of them.³ Be that as it may (and the facts of those matters are far from settled), I shall here be talking instead about terms of "target efficiency." That notion of efficiency is concerned strictly with assessing how well a program achieves its own goals, with how many "hits" it scores on its own self-defined "targets."

According to the conventional wisdom, overly generous social benefits are inefficient in that sense, as well. Paying the same sum to everyone, willy-nilly, seems inevitably to deliver less cash to those who really need it than would distributing the same overall amount of money through more tightly targeted payments conditioned on tests of people's needs and of assets. Of course there are many things to be said against means-testing social benefits; and many of those familiar objections touch directly or indirectly upon the target efficiency of the practice.⁴ There is, however, another way in which attempting to target social benefits paradoxically actually undermines the target inefficiency of those programs.

The essential point is just that targeting always seems to end up working through surrogate indicators. Family size is taken as an indicator of how much food and clothing a household needs. Income is an indicator of how many calories will appear on the dinner table. Marital (or household) status is an indicator the potential spending power of those who themselves earn no wage. And so on.

All those indicators of needs, assets or means – and still more, all of the social categories built into more explicitly "categorical" programs of social security – serve merely as surrogates for what morally matters for purposes of our social policy. Inevitably, upon closer inspection we find that many of macro- and meso-sociological assumptions embodied in present social policies are no longer valid (and maybe they never were).⁵ Perhaps once it was safe for policy-makers to assume income-sharing on an equitable basis within families; they can no longer. Perhaps once they could assume that most families

³ Blaug 1963. Murray 1984. Okun 1975. Cf. Goodin 1988, chap. 8; Barr 1987, chaps. 4 and 5; 1989; Van Parijs 1990.

⁴ Means tests are stigmatizing; and in part because of that they reduce take-up of social benefits, often among those who most need them (Deacon and Bradshaw 1983; Goodin 1988, chap. 7). Furthermore, the nonneedy often find ways of masquerading so as to pass even the harshest tests of needs and means (Goodin and Le Grand 1987, chap. 6).

⁵ Titmuss ([1967]1987b, chap. 7) argues similarly, in his classic paper on "Universal and Selective Social Services."

would have at least one breadwinner permanently employed at least most of the time; again, they can no longer. And so on down the list.

Insofar as the sociological facts of the matter have changed, that would constitute a strong case against social policies targeting benefits on the basis of those false assumptions. Some of those new sociological facts – primarily facts about infirm work-force attachment on the part of much of the population – might in themselves argue directly for the distinctively “nonproductivist” mode of social provision embodied in basic income schemes.⁶

Here I propose to develop an even stronger argument for basic income by building more obliquely upon those points about changing social dynamics. The aim is to produce an argument that is not contingent (or anyway much less contingent) upon the truth of any particular set of sociological propositions. Rather than arguing that the basic-income strategy is somehow more accurate in its underlying assumptions about social reality, I shall be arguing that it is simply more *noncommittal* – it simply makes fewer assumptions. In so doing, it manages to be less prone to sociological error and less vulnerable to social change than are alternative modes of social security provision. That, in itself, contributes heavily to the target efficiency of such schemes in a world in which sociological facts are uncertain, highly variable and, in any case, constantly changing.

1. THE PARTICULAR PRESUMPTIONS OF PRESENT POLICIES

Let us start, then, by examining the various sociological presumptions built into present social security policies. My strategy here will be to focus tightly upon particular programs, their administrative arrangements and their benefit structures, in search of the substantive sociological presumptions that are embedded in them.

Notice, first, that across the world – in both Bismarck and Beveridge countries alike⁷ – administrative arrangements for vesting social security pensions presuppose something very much like “full employment.” In virtually all those countries, the old-age pension is a contributory scheme, with the predominant mode of vesting entitlements being through employment-based contributions (typically, matching contributions from employers and employees) to the social security fund on the employee’s behalf.⁸

Those contributions typically serve to vest pension rights not only

⁶ Offe 1992.

⁷ That is, in at least two of Esping-Andersen’s (1990) “three worlds of welfare capitalism.”

⁸ U.S. DHEW 1978. Palme 1990.

for employees but also for their dependents (typically, spouses and dependent children of tender years). So, strictly speaking, this arrangement for vesting pension rights does not quite presuppose full employment in the sense of literally everyone of working age being in paid employment. But it does in effect presuppose that everyone in society is a member of some household unit, and that all heads of households are in paid employment. It presupposes, that is to say, that everyone is either in work or in a stable, long-term liaison with someone who is.

Now, in a way this presumption constitutes little more than an administrative convenience. Like "pay as you earn" arrangements for withholding taxes from wages (and for withholding taxes at source more generally), withholding social security contributions from wages is undeniably convenient. The state is more certain of getting its money that way; and individuals are thereby relieved of the burden of having to find large sums of money all at once, quarterly or at the end of every year.

If employment-based contributions to a pension fund were no more than an administrative convenience, then there should be other logically possible ways of vesting one's old-age pension entitlements. Given the administrative convenience of such payroll withholding arrangements, that is of course how pension entitlements would accrue for those actually in employment. But for those who are unemployed or nonstandardly employed, other sorts of arrangements could in principle be made. The self-employed could contribute – or even be required to contribute – to the pension fund to vest their own pension rights. For those who are persistently unemployed or literally unemployable, the contribution requirement could be waived altogether; they could be allowed to draw a pension without having made any contributions. Or, if we insisted upon retaining the fiction that contributions are needed to vest one's pension rights, the state agency responsible for the relief of the unemployed and unemployable might be required to make periodic payments into the pension fund on each client's behalf.

It is nonetheless striking how infrequently and imperfectly such possibilities are actually taken up. In the United States, social security pensions are simply unavailable to those without the requisite employment-based contribution history. The permanently unemployed and unattached may well be able to claim other means-tested social benefits, but they cannot claim old-age (social security) pensions. And even where those who have made no contributions are nonetheless allowed to draw the old-age pension, those arrangements are inevitably awkward for all concerned. Insofar as social insurance schemes are run principally on the basis of workplace contributions, it seems

that the unemployed will always be second-class participants in such schemes.

A second central presupposition of present social security policy is that people for the most part live in stable family units where income is shared equitably, if not necessarily equally.⁹ Some such presupposition about family life is revealed in the way, just mentioned, in which dependents can claim Social Security benefits on the strength of workplace contributions made by the head of the household. But it also figures centrally in other aspects of social welfare policy.

Consider, for example, the way in which means tests characteristically work to assess the family's need for social assistance. The test is not on the means of each individual member but is rather on the means of the household as a whole. The presumption, obviously, is that households pool all their available resources, at least in times of crisis. Or consider, again, the way in which U.S. authorities used to calculate Aid to Families with Dependent Children. For many years, in figuring the amount of aid due to a woman they would simply assume that any child-support payments which courts have ordered would actually be paid by the absent father, even though defaulting fathers had typically moved outside the jurisdiction of the court in question. Or consider, yet again, the way in which British authorities operate the means-tested grant to students in higher education. They calculate how much each student's parents should be able to afford to pay and they then award the student a grant representing the difference between that amount and the full grant, on the assumption (which is merely an assumption, which state authorities make no attempt to enforce) that the student's parents will actually contribute whatever they have been putatively assessed.

Again, these presumptions are not strictly necessary to the policies in question. If we doubted that husbands shared income equitably with wives (and the evidence suggests that they do not)¹⁰ we could assess each on an individual basis rather than all on a family basis for all sorts of purposes – taxes, means-testing of social benefits, and so on. If we doubted that child support payments are necessarily forthcoming, whatever the courts may order, we could pay the single parent the whole AFDC stipend on condition that she transfer child-support claims to the state, who would then press the claim against the delinquent father, in turn.¹¹ If we doubted that parents will voluntarily pay as much toward their child's education as we think they should, we could pay the child the full grant and collect the putative

⁹ The American case is well discussed by tenBroek (1954).

¹⁰ M. Edwards 1981. Piachaud 1982. Pahl 1983.

¹¹ As under the U.S. Social Security Amendments of 1974 (Mondale, Bentsen and Ribicoff 1974).

"parental contribution" through a special tax levied against that child's parents. None of this is impossible, and much of it has already happened to some extent in one place or another. But while no particular presumption is necessary, some presumption or another is again apparently necessary in all such schemes of categorical, targeted, means-tested assistance.

Note, finally, a third set of related presumptions built into recent social security reforms instigated by the New Right. Those reforms are couched in terms of catch-phrases like "self-reliance," "community care" and "mediating structures." At root, all involve transferring large portions of the state's welfare functions to families, friends and voluntary community charities.¹² And, more often than not, they involve transferring those responsibilities to female caregivers within those groups.¹³

Of course, even the New Right envisages a system of residual safety-net programs of social assistance to cater to those who excite little sympathy among family, friends or charitable institutions. But like all safety-net programs, being a residual program, that one will be deliberately spare and underfunded. The expectation is that most cases, and the unspoken implication is that almost all "really deserving cases," will be adequately catered to elsewhere.

Much might be said against such schemes. The questionable morality of imposing on already overworked and underpaid female caregivers in this way makes such schemes an anathema to feminists – and rightly so. But quite apart from all that, there is also the simple empirical question of whether such schemes could possibly hope to work as envisaged given a society such as ours now is.

Underlying all such schemes for devolving the state's welfare function onto private parties – and for dramatically running down public provision in consequence – is the basic assumption that most people in need actually have someone to whom they can turn in this way. There is much in the sociological literature, though, to suggest that social support networks are much less dense than they once were. There is much to suggest that trend will continue, with the increasing shift of females (the traditional care-givers) out of the home and into the labor market. In short, recent New Right social welfare reforms presuppose a sociological reality that is rapidly waning and is unlikely to be re-created.¹⁴

¹² These proposals from New Right writers like Berger and Neuhaus (1977), Murray (1984) and Novack (1987) won the qualified support even from the OECD (1981). For analyses of how such schemes work in practice, see Johnson 1987; 1990; for an ethical critique, see Goodin 1988, chap. 12.

¹³ Land 1978. Braithwaite 1990.

¹⁴ d'Abbs 1982. Blumer 1987.

My problem is not so much with the particular assumptions currently embodied in these or any other particular social security policies, though. It is not as if those assumptions were somehow inherent in the very fabric of present social security policy. Alterations of various sorts could be, and in many places have been, incorporated in such policies without any undue strain to their fundamental fabric. The point is just that, where social policy does incorporate any specific assumptions, we must face that fact that they will inevitably need to be altered in light of changing social knowledge and changing social circumstances. And, as I shall argue in Section III, those alterations will inevitably lag somewhat behind (and possibly very far behind) the changes occasioning them. Hence, policies will inevitably fail to fit social reality for some (possibly considerable) time, insofar as they are predicated on any particular sociological assumptions, in this way.

II. THE ROLE OF PRESUMPTIONS IN THE PRESENT POLICY STYLE

Let us step back, now, from particular policies and consider the role of presumptions in relation to whole "styles" of social security policy, more generically considered. The particular presumptions embodied in any particular policy may be peculiar to it alone, and can be altered easily enough merely by changing the policy. But presumptions of one sort or another are endemic in the present style of social security policy per se. If - as will turn out to be the case in Section III - our objection is to the use of presumptions as such, rather than to any particular presumptions that we find flawed, then to meet the challenge a whole different style of policy rather than merely a different particular policy will be required.

There are, of course, many ways of identifying the "present style of social security policy." Its central feature can, for present purposes, be said to be the fundamental tendency to distinguish between two tiers of cases, for purposes of social security policy. On the top tier, relatively more generous sums are paid to claimants relatively more automatically. On the bottom tier, more niggardly sums are paid to claimants, subject to more discretion on the part of the administrator responsible for their case.

These two tiers can be variously characterized. One characterization distinguishes between programs of categorical and of general assistance, another between contributory and noncontributory programs, still another between social insurance and social assistance. Those various distinctions are largely overlapping, of course, so it might not much matter which characterization we choose. What is important to note, however, is that whatever distinction we employ for separating

the favored top tier of social programs from the disfavored bottom tier must necessarily embody presumptions of some sort.

Programs of categorical assistance pick out for especially favorable treatment certain classes of people: those too old to work, or the blind, or the victims of occupational injuries or criminal assaults, or whatever. But questions will then always arise as to why we should privilege those – and only those – categories of people in our social policy.¹⁵ If we are going to privilege those too old to work, why not those too young, or those unable to take paid employment because they are burdened with the responsibilities of caring for the old or young at home?¹⁶ If we are going to privilege those who are blind, why not those who are deaf and dumb, or those with debilitating physical or mental impairments?¹⁷ If we are going to run a criminal injuries compensation scheme, why not similarly compensate out of public coffers victims of mere accidents; and if we are going to help those injured in accidents, why not help similarly victims of diseases as well?¹⁸

Of course, none of those particular choices is rooted in the nature of the program itself. We may well be led by such logic toward an ever expanding program of categorical assistance, embracing more and more cases with broader and broader categories. My point is not that that is impossible. My point is, instead, just that making any sharp distinctions between programs of categorical and general assistance in this way necessarily entails presumptions.

First, it presumes that the right categories have been identified: that persons in the categories picked out for better treatment are importantly different from everyone else and that they are, at least on average, more qualified for that better treatment. Perhaps some blind people find it easier holding down a job than do some neurotics, but unless the converse were characteristically the case there would be no grounds for including the blind but excluding the neurotic within the set of categorically privileged claimants.

Second, the practice of categorical assistance presumes that the right treatment differential has been established. That is to say, we presuppose that the extent of the differential treatment accorded to those in the categories selected for especially favorable treatment corresponds to the actual differences – again, at least on average – between them and everyone else, along whatever dimension it is that justifies that differential treatment in the first place. Thus, for example,

¹⁵ The broader issues involved here are well cataloged by Tussman and tenBroek 1949.

¹⁶ Titmuss [1955] 1987a.

¹⁷ D. Stone 1984.

¹⁸ Atiyah 1975, chap. 20. See further Chapter 11 in the present volume.

in making short-term sickness benefits half again larger than long-term disability benefits, the presumption must be that people are usually able to rearrange their economic affairs so as to reduce their financial commitments by something like 50 percent during the six months they spend on the short-term, higher benefit before they are shifted over to the lower long-term disability pension.¹⁹

A similar story can be told about all the other ways of characterizing our two-tier social security system. With contributory or social insurance schemes, who belongs on which tier seems clearer, at least at first blush. Notionally, the distinction is simply between those who have contributed (i.e., who have paid an insurance premium) and those who have not. But matters are never so simple. By that rule, we should refuse to pay contributory benefits to anyone who has not himself contributed – including, for example, those already of pensionable age when the contributory old-age pension was established or, for another example, those who have not themselves contributed but who are dependents of deceased others who have. In practice, we always have extended the benefits of contributory social security schemes to such persons right from the start. If we are going to relax the contributory logic that far, though, why no further?

There must be some presumptions at work here – presumptions about fairness or the structure of family life – that justify those but rule out other apparently similar extensions of contributory benefits to noncontributing parties. Presumably, for example, the ground for pensioning those who were already old at the inception of the old-age pension scheme was that it would be unfair, somehow, to deprive them of the benefits of a scheme to which they would happily have joined, if they had only had the chance. But by that logic, we should similarly include the persistently unemployed and the unemployable. They too, it could equally plausibly be argued, would have been happy to join the scheme had they only been given the chance of a job.

Here again, two-tier policies presume, by their nature, that we have identified the right principles to use in placing people on one tier or the other. We presume that contributors (or those we are, for reasons other than contribution, prepared to treat as contributors) are importantly different – at least on average – from everyone else, in some morally important respect. We presume, furthermore, that the difference in the treatment we accord them is proportional to those differences, again at least on average.²⁰

¹⁹ See Chapter 12 and sources cited therein.

²⁰ The latter presumption wears particularly thin in social insurance schemes involving compulsory risk-pooling, where better-than-average risks are in effect

The point, here even more emphatically than before in Section I, is not that any particular presumption is necessarily embodied in social security policy organized on a categorical, or contributory, or social-insurance basis. If we decide one presumption is in error, we can easily enough shift to some other on which to base our categorical distinctions or the treatment differentials accompanying those distinctions. The point is simply that we *do* have to shift to some other presumption. If we are to sustain that sort of two-tier social security program at all, we simply cannot do without presumptions altogether within such schemes. We need some basis for distinguishing between cases on the two tiers, and with such bases for distinction inevitably come presumptions of one sort or another.

III. PROBLEMS OF PRESUMPTIONS

The basic problems with presumptions are plain for all to see. They are vulnerable to factual error, and they are vulnerable to social change. American social security policy presupposes, in its administrative structure, that most households have a breadwinner who is employed most of the time. That certainly was not true when the Social Security Act was passed into law during the Great Depression; it certainly is not true of a small but significant minority of the American households, even today. The structure of the Social Security Act similarly presupposes that families stick together through time, so that social security entitlements built up for the couple as a whole during the breadwinner's working years will actually cover the same pair in their old age. Such stability may well have characterized American family life in the 1930s but it certainly does not today.

Of course, it is perfectly true that social security policy always contains a residual form of assistance extended to people who are in need, whether or not they fit into any preordained categories. Those not entitled to social insurance can always fall back upon social assistance. Inevitably, however, such residual, catchall policy categories are treated less favorably by policy-makers. They inevitably are funded less generously, and they inevitably carry more social stigma than do their mainstream policy counterparts. So it is simply not acceptable to discount the importance of any errors we might make in our policy by saying that there is a residual safety-net policy to catch any mistakes we might have made. Where we place the emphasis – what we take as the paradigm cases and what we take as the peripheral ones – clearly does matter.

legally compelled to cross-subsidize worse-than-average risks (Titmuss 1968, pp. 173-87; Goodin 1988, chap. 6).

That points to the third way in which social security policy is vulnerable to presumptions. It is vulnerable to errors in emphasis, treating as exceptions or residuals what are truly standard and paradigmatic cases of the social problems a policy is supposed to be meeting. When setting up a program, we try to tailor it to what we suppose to be the standard case. If we are mistaken in what we suppose to be the standard case, or if the standard case changes, our policy is not the altogether fitting response to the standard case that we suppose it to be. Administrators will have increasingly to twist and turn to assimilate rather different cases, now standard, within the structure of a policy designed with a different paradigm case in view. The fit will never be perfect, and the injustice will therefore always be palpable.

None of this would matter, of course, if social policy were instantly alterable once we realized our errors. But of course it is not. Indeed, a veritable presumption in favor of the status quo sometimes seems to be at work – especially in purely administrative settings – with those advocating change being required to carry not just a rhetorical but almost a quasi-legal burden of proof that a change is needed. How seriously we should take such formalistic talk about burdens of proof and presumptions outside such settings is, perhaps, unclear. Still, in purely practical political terms those advocating change must inevitably be the ones to bear the burden of mobilizing a coalition. The law stands as written until it is rewritten, and change inevitably requires further legislation or administrative action in a way that sticking with the status quo does not. If the errors are palpable and their consequences severe, the change will no doubt eventually come. But given the costs of organizing changes, doubtless it will not come immediately, and serious injustices will occur in the meanwhile.

In arguing against presumptions in this way, we must not overlook possible arguments in their favor. Some presumptions have little to be said in their favor. They are no more than descriptive presumptions – presumptions about matters of fact. When those turn out to be in error, there is no good reason not to change them straightaway. But presumptions can be prescriptive as well as descriptive. In that latter case, there may well be good grounds for persisting with presumptions we know to be descriptively inaccurate.

In many of the examples cited earlier, it may well be prescriptive rather than descriptive presumptions that are at work. We presume for purposes of social policy that families will share income equitably, that families will stick together, and that people will respect court orders to support their children. But perhaps we are under no illusions about the empirical unrealism of those propositions, in all too many cases. Perhaps we build those presumptions into social policy

nonetheless because we want to try to alter social reality in those respects, and we think that building those presumptions into social policy will somehow help to do so.

The goals embodied in all those prescriptive presumptions are undeniably worthy. The question is just whether the best way to pursue those goals is by building them into social programs in the form of presumptions. Usually it is not. If we want to enforce court orders against nonsupporting parents, surely we would be more successful in using the ordinary legal mechanisms for enforcing court orders – attaching their wages, requiring surety bonds or, ultimately, incarcerating them for contempt of court.¹¹ If we want to make sure that families support their children through higher education, surely it is better to tax them for the sum we presume they should be able to pay. If we want to ensure that husbands and wives share family resources equitably, surely it is better to treat them as individuals rather than a family unit for tax and social security purposes, so each gets the same sum rather than the family as a whole getting a lump sum to be divided as it sees fit between its members.¹²

Simply setting up policy in such a way as to presume that these desirable behaviors occur is a less direct, and for that reason less effective, way of guaranteeing those results. All too often, it is no use at all, because when the policy fails to achieve its goals the penalty for that failure falls on the wrong party – the starving student or the undervalued spouse or whomever.

IV. BASIC INCOME AS A MINIMALLY PRESUMPTUOUS STRATEGY

For all those reasons, it is desirable that our social policies should rely as little as possible upon presumptions. The basic income strategy can be commended, against that background, as the minimally presumptuous form of social security policy.

It achieves that honor, of course, by also being the least discriminating form of social security policy. Whereas other forms of social

¹¹ Chambers 1979. Withholding payments from wages increases the ratio of child-support dollars paid-to-owed from some 53 percent to some 70 percent; but that still leaves a stubbornly recalcitrant 30 percent of support awards unpaid, even after the most draconian treatment we can seriously contemplate instituting (Garfinkel and Klawitter 1990).

¹² Although that is the larger implication of the recent work on family income-sharing (M. Edwards 1981; Piachaud 1982; Pahl 1983), in practice most of those studies focus on more modest policy goals such as retaining the child benefit, which is in practice collected by the (typically female) partner who is not in paid employment and who is in consequence typically disadvantaged in the intrafamily income carve-up.

security policy attempt to distinguish between more and less deserving categories of claimants, the basic-income strategy attempts no such distinctions. It simply gives everyone in the society the same basic income.

There are obvious costs and disadvantages to being so indiscriminate in our policy, of course. But there are also advantages. By attempting fewer discriminations, the basic-income strategy also manages to get by with fewer presumptions. In other more discriminating two-tier social security strategies, it is crucial to specify our categories correctly and to set the treatment differentials between them correctly. Sociological presumptions which may or may not (now) be empirically warranted enter as we try to establish any such distinctions. An advantage of a policy attempting no such distinctions – whatever its other disadvantages – is that at least no such precarious presumptions get smuggled into social policy.

Thus, for example, programs of categorical assistance attempting to aid only those genuinely unable to work need to draw sharp lines between one sort of disability and another, or one cause of impairment and another. In so doing, those programs make various sociological assumptions – assumptions about the nature of paid work and about what is and is not compatible with performing it in our society. Those assumptions are naturally prone to sociological error and to social change, in the ways just described.

A basic-income approach runs no such risks. The reason is quite simply that it takes no notice of why a person's income is low – or, indeed, of how high or low his income is. Whether a person is unable to work or merely unwilling to do so is of no consequence, in terms of a basic-income approach to social security. Such an approach merely arranges, without further questions, to bring everyone's income at least up to some stipulated social minimum.

That example was chosen deliberately. It shows the basic-income strategy in what would seem to be its worst light. Or at least I presume that most people would intuitively scorn a policy that proposed to pay identical social benefits to the severely disabled and to the merely idle. Contrary to our immediate intuitive reactions, however, I now want to argue that that is the right thing to do.

Perhaps we should, if only we could, pay benefits to the deserving and only the deserving. But those categories are not themselves clear-cut. The “undeserving” might look a lot more deserving if we took into account alternative forms of socially useful activity (in caring for young children or aged relatives, e.g.) performed outside the labor market. Or, again, we might count as “deserving” all those who would work if provided good jobs at decent wages – thus shifting the onus from unemployed individuals to the society that fails to provide

them decent employment opportunities.³¹ And all of this in a way merely betokens what are truly massive conceptual problems with notions of moral desert, problems which prove particularly vexing in these realms.³²

Even assuming the concepts of “deserving” and “undeserving” are clearer than they actually are in these realms, we would face the problem of applying them to particular cases. There is, as discussed earlier, always a serious prospect of sociological error and of social change in these matters. Those risks vitiate, in turn, any categorical distinctions with which we might try to capture those two concepts in our policies. The risk of error is always with us, and the question is simply on which side we should prefer to err. Is it more important to deny benefits to the undeserving, even if that means denying it to some of the deserving, too? Or is it more important to make sure that all the deserving get the benefit, even if that means giving it to some of the undeserving, too?

The answer is, of course, sensitive to actual numbers. If getting the benefit to the last dozen deserving cases means giving it to thousands of undeserving cases, we might decide that on balance it is better to deny it to those last few deserving cases. Conversely, if weeding out the last dozen spongers means denying the benefit to thousands of deserving cases, we might decide it is on balance better to accept that a few people will get undeserved rewards.³³

The calculation upon which the case for an indiscriminate basic-income approach to social security policy rests is just this. The harm done by deserving cases being denied benefit by errors that accompany a more discriminating policy are, it is judged, worse than is that done by undeserving cases being granted benefits by a less discriminating policy. In terms of the prejudicial example with which I introduced this discussion, it is judged on balance to be worse to take the chance that errors in categorical social security policies will deny benefits to those genuinely unable to work than it is simply to swallow the fact that some idle louts will get a basic income they do not deserve. Put that way, the repugnant conclusion is surely more appealing.

V. PERSISTING PRESUMPTIONS

Although the basic-income approach is the minimally presumptuous social security policy, it is not without presumptions altogether. There

³¹ As Claus Offe has usefully reminded me.

³² Goodin 1988, chap. 10.

³³ Goodin 1985b. Atkinson (1990) has powerfully extended this analysis, with particular reference to unemployment benefits in the United Kingdom.

are some presumptions – some of which are genuinely problematic ones – inherent in an income-support scheme of any sort, one that relies on notions of a basic income included among all the others.

Social security programs generically are income-transfer schemes, aimed at providing compensation to certain categories of persons for certain kinds of losses. The basic-income approach just generalizes this strategy, providing cash to people unconditionally, without further inquiries about what categories they might fall into or what sorts of losses they might have suffered. But it still shares the same root presumption underlying all social security policy – namely, that whatever it is that might ail people, an infusion of additional cash will cure it.

There is much of which that is true. But there is just as much of which it is not. Suppose someone has suffered an injury resulting in the loss of income – the classic sort of case traditional social security (workmen's compensation) policies are designed to remedy. Now, if what people care about in income is just the money, then an infusion of money from one source can easily enough substitute for money lost from another. There, cash transfers can compensate perfectly well.

But even with so easy a case as that, matters are often more complicated. People sometimes (to some extent, probably always) care about the source of the income as well as about the money per se. Earned income is a source of pride in a way that unearned income is not. And, as has been shown in Chapter 12, matters get more complicated still when the loss in question is not just monetary. Of course we can give people money to compensate for the loss of an arm; but the wad of bank notes does not look or perform at all like the arm used to do.

Now, social security schemes in general, the basic income approach included, are all mechanisms of income support. They all work through compensation for losses. The presupposition that they all share is that losses can be fully compensated. And, insofar as they are income-transfer schemes, they further presuppose that the losses can be compensated by monetary payments. Insofar as certain losses cannot be compensated, or cannot be compensated by monetary transfers, such presumptions are simply unwarranted.

The implication, as I see it, is merely that there are limits to what social security policy – however organized – can do for us. For certain sorts of irremediable losses, we must pursue a policy predicated on goals of prevention rather than of ex post compensation. Regulations to prevent workplace injuries are better than disability pensions for those suffering such injuries, for example, precisely because a disability pension provides only lost wages whereas the injured worker has lost an arm as well as the lost wages for which the pension might

compensate him.²⁶ And insofar as work – and earnings more generally – is a source of pride as well as of money, a full employment policy preventing loss of pride-inducing income would be preferable to a universal grant, unrelated to work, providing the same monetary sum.²⁷

VI. CONCLUSION: EFFICIENCY IN ITS PLACE

In this chapter, I have been making principally two claims. One is that proposals for an unconditional basic income for all constitute a minimally presumptuous social policy. The other claim is that policies which are minimally presumptuous in that way are likely to be maximally target-efficient, at least in a world where sociological facts upon which more categorically based programs might be conditioned are uncertain, highly variable and, in any case, constantly changing.

Arguing for basic income in efficiency terms in that way is undeniably cheeky. It meets and beats critics of social benefits in general, and of unconditional ones in particular, on their own chosen ground. In a context in which apparently the worst thing that anyone can say against such programs is that they are inefficient, the best thing to say on their behalf is that, quite the contrary, such programs may well be more efficient than any of the more draconian alternatives that their critics would prefer.

In closing, though, it is important to remember that we do not value efficiency in and of itself. Instead it is valued merely as a means. Whatever our other goals may be, we want to pursue them as efficiently as possible merely so that we may achieve as many of those other goals as possible. Efficiency allows us to achieve more of those other valued goals. That is its value – no more, no less. Its own value of efficiency is thus wholly derivative from the value of those other goals. Efficiency as such is of no independent moral importance to us.²⁸ So at root the reason we should cherish the target efficiency of basic income strategies is simply that that guarantees that we will, through them, be able to relieve human suffering as best we can.²⁹

²⁶ As argued in Chapter 12.

²⁷ The same conclusion – that we should strive to get the primary distribution right rather than counting on a secondary posttransfer distribution to provide a remedy – is urged on other grounds in Goodin and Le Grand 1987, pp. 222–5. A model might be the Australian “workers welfare state,” where social security was traditionally pursued more through wage policy than transfer payments (Castles 1985).

²⁸ Goodin 1988, pp. 245–56. Le Grand 1990.

²⁹ I am grateful to Claus Offe and Philippe Van Parijs for comments on an earlier version of this chapter.

Relative needs

What conventionally follows in public policy terms from recognizing something as an unmet need is utterly uncontentious. It is standardly thought to be indisputably better for people to be provided with more of what they need, up to the point that they need no more of it. What follows from the proposition that there is a need for more housing? That we should build more of it, obviously. What follows from the proposition that there is a need for more education? That we should supply more of it, surely. What could be more straightforward?

It is the theme of this chapter that, perhaps surprisingly, maximizing the supply of needed resources is not always the right response to unmet need. At least for certain important classes of needed resources – and perhaps for most classes of needed resources, at least across certain portions of their ranges – the best way of meeting unmet needs may entail decreasing rather than increasing supply. More specifically, the best way of satisfying unmet needs may not be to cause those who are relatively more needy to have more of the needed resource, but rather to cause others to have less of it.¹

This chapter starts from two commonplaces about needs, their priority and their relativity. It proceeds to develop a paradox, arising principally out of the latter attribute but deriving much of its sting from the first. The counterintuitive conclusion just indicated is then teased out of that paradox, and its implications for public policy in various areas are explored.

¹ I take my text here from Rousseau ([1762]/1911, p. 45): "Misery consists, not in the lack of things, but in the needs which they impose... 'Great needs,' said Favorin, 'spring from great wealth; and often the best way of getting what we want is to get rid of what we have.'"

I. TWO COMMONPLACES

It is often said that, in any conflict between satisfying people's desires and meeting people's needs, the latter is to take priority. Intrapersonally, that proposition constitutes the core of the case for paternalism. It tells us to serve a person's needs, whether he or she wants them satisfied or not. Interpersonally, that proposition guides the allocation of social resources, giving priority to the satisfaction of one person's needs over another's "mere desires."

Whether it is needs as such that deserve such high priority treatment, or whether it is merely something that correlates with needs, is an open question. But that correlation – if that is all that it is – is typically thought to be strong enough for claims of needs to exert a strong moral pull on us. Exactly how strong a pull is another open question. But the priority to be accorded to needs over wants is typically said to be very strong, and is often said to be virtually (if rarely literally) absolute.¹ That is the first commonplace with which I shall here be conjuring.

The second commonplace is that needs can be relative as well as absolute. There are in fact many ways in which, and many things to which, needs might be relative. Here I shall focus on one of them in particular, so let me just mention the various others to set them aside. Most fundamentally, needs are of course relative to the goals for which the resource in question is needed. Needs are inherently instrumental in that respect.² Beyond that, needs are relative to the society (time, place and social circumstance) in which you find yourself. You have no need for a fur coat in Havana, for fissionable materials in ancient Greece, for mainframe computers in Javanese rice fields, for hunting lodges if you are in the English working class or for football tickets if you are of the landed gentry.³

Here, however, I shall be focusing instead on a deeper sense in which needs might be relative. My point is not just that needs are essentially conventional, or in some other way relative to time, place and social circumstance.⁴ That is true, too. But that is not the end of

¹ Goodin 1988, chap. 2. Cf. Braybrooke 1987; Frankfurt 1984.

² Barry 1965, pp. 47–9. Braybrooke 1987, chap. 2. Wiggins 1985, pp. 154–5. Goodin 1988, pp. 29–32.

³ The latter point, originally Townsend's (1962, p. 219), forms the basis of Pichaud's (1981) powerful critique of his later work in which Townsend (1979) seems to have forgotten it.

⁴ "By necessities I understand . . . whatever the custom of the country renders it indecent for creditable people, even of the lowest order, to be without," in Adam Smith's phrase upon which subsequent work on "relative poverty" builds (Townsend 1962, p. 219, quoting Smith; Townsend 1979).

the matter. What in particular you need does not just depend on the society in which you are living.

My focus here will instead be on the way in which, within a single society, what one person needs may also be relative. My needs may be relative, not only to the society in which I live, but also to what others in my society already have. One person's needs-fulfillment may be a function (positively, or more typically negatively) of other people's needs-fulfillment in the same society. What resources one person needs in order to satisfy the goal of meeting certain needs may be relative to the resources available to others in the same society for meeting their similar goals.⁶

The former style of social relativization of needs is a model of "relativism." It relativizes needs to societies. What you need depends on the society (its objective circumstances and its extant values) in which you find yourself. The latter and deeper style of social relativization of needs is genuinely a model of "relativities." It relativizes your needs to the needs-satisfactions of other members of your society. Your needs-satisfaction, on that latter model, depends not just on the society in which you are living but also on how well you, compared to everyone else in that society, are doing toward meeting the standard that that society sets as the benchmark of needs-satisfaction. Henceforth when talking of "relative needs" I shall mean the term in that latter, stronger sense.

There are clearly certain needs that are, by that standard, not relative at all. They are completely fixed by nature and wholly independent of social context. The minimum caloric intake required to sustain a human body of given size and weight at a given level of physical exertion in a given climate might be one such example.⁷ At the very least, we can say with confidence that how much sustenance you need to exert a certain amount of force on the natural world is independent of the caloric intake of others around you.

Other needs, though, are more dependent on social context. The clearest examples, perhaps, are those arising in connection with Peter Townsend's discussion of "relative poverty."⁸ Whether you need access to a television set to participate in social conversations in your society depends on what proportion of the population around you

⁶ There are what Sir James Steuart ([1767] 1966, bk. 2, chap. 21) dubbed "political necessities" required for pursuit of "rank in society," distinguished from "physical necessities" required for mere survival.

⁷ Dasgupta and Ray (1986-7) point out how dietary deficiencies handicap the malnourished in labor market competition - in that sense, your need for food really would be relative to others' consumption. See more generally Streeten 1981 and Townsend 1962, pp. 216-18.

⁸ Townsend 1954; 1962; 1979.

has access to one, and on how much social conversation is dominated by last night's programs in consequence. How many calories you need might even depend on what society expects you to do by way of expending calories, which in turn depends on how many calories others around you have available to expend. If they are all playing tennis, you need to have the strength to join in the game if you are not to be excluded from that aspect of social life altogether.

In all such practical illustrations of the social relativity of needs, it is almost inevitably unclear which sense of relativity – strong or weak – is at work.⁹ Sometimes the point of an example seems just to be that needs are conventional, and relative to time and place and cultural location in that sense. (Townsend's assessment of Britons as relatively poor if they cannot afford to eat roast beef on Sunday seems a clear example of that.)¹⁰ At other times, though, the point does genuinely seem to be that your needs depend not only on what local convention requires of you but also on how well everyone else in your society satisfies those requirements. Your own needs-fulfillment really is then dependent on, and relative to, theirs. What you need truly depends on what they already have; the more needed resources they have, the more of them you need.

II. THE PARADOX

If needs are relative in that strong sense, then there are in principle two quite distinct ways of meeting needs. One is to increase the supply of needed resources to the relatively more needy; the other is to reduce the supply of needed resources to the relatively less needy. Or, of course, we might pursue the mixed strategy of doing both at the same time.

Now, there is nothing paradoxical about the first option. It is not in the least odd to suggest that we reduce aggregate neediness in the population by giving more to the relatively more needy. There is nothing paradoxical in the suggestion that aggregate neediness could be reduced by distributing any extra resources disproportionately in favor of the relatively more needy, for example. Nor is there anything paradoxical in suggesting that we pursue that goal by redistributing needed resources away from the relatively less needy and toward the relatively more needy. Such strategies, insofar as they are designed to

⁹ This is in part because the strong sense actually implies the weak. If what you need is relative to what others around you have, for individuals one by one, then it follows that what you need will also be relative to some society-specific aggregate (mean consumption, or some such).

¹⁰ Townsend 1979.

give the relatively more needy people relatively more needed resources, are intuitively wholly accessible and perfectly appealing.

Notice, however, that the second option – equally eligible, on the logic of meeting relative needs – does not stop with anything nearly so unexceptionable as that. It is concerned primarily that we should take needed resources *away* from the relatively less needy people in the population. What we then do with those resources is, under that second strategy, strictly speaking of no concern. To be sure, we would make more of an impact on the reduction of relative needs if we were to conjoin the second strategy with the first, both taking from the relatively rich and giving to the relatively poor. But in the unalloyed version of the second strategy, that is not strictly necessary. The desired effect could be achieved perfectly well simply by taking resources from the relatively less needy, and then throwing them into the sea. That would be a perfectly good, albeit perhaps nonideal, way of reducing the relative disparity between the two groups – and in so doing, reducing the relatively more needy group's unmet (relative) needs.

The logic of that inference is impeccable. Yet there is something undeniably paradoxical about meeting needs by wantonly destroying needed resources. There is something unquestionably odd about recommending impoverishment as a strategy for needs-satisfaction.

Of course, the general structure of the problem is familiar from other contexts. But the standard ways we have for solving such problems in those other contexts are generally unavailable to us here. It is often said, for example, in discussions of equality that if all we care about is equality per se then leveling down is as good as leveling up. Throwing the treasures of the rich into the sea is a perfectly good way of producing equality, too. If enough of the rich's treasures are tossed away, all would be equal – albeit equally poor.

The standard response, in the case of equality, is that equality is *not* all we care about, nor does equality even enjoy any particularly strong priority over everything else that we also care about. Thus, for example, we care about minimizing poverty or maximizing the average standard of living as well as about minimizing inequality. And it is those other, competing, compelling goals that make "leveling down" an unacceptable strategy of equality.

The same response, however, works less well in the context of arguments about relative needs. Needs claims have quite high (if not quite absolute) priority over all other claims. The stronger we think that that priority is, the more relaxed we should be about sacrificing other goals to them, through a socially wasteful process of leveling down to maximize relative needs-satisfactions. Of course, which strategy we should ultimately prefer depends on what alternative mech-

anisms are available for satisfying relative needs, and on how effective they happen to be. But it is perfectly possible that the goal may be of sufficiently high priority, and the alternative means of pursuing it sufficiently unpromising, that the strategy of wantonly destroying the resources of the rich, Pol Pot style, might turn out to be the socially preferred strategy.

Here, then, the paradox resists the standard solvents. It is not just that some lower priority social goal (efficiency, or whatever) is suffering unnecessarily when we throw some needed resources of the relatively less needy into the sea, and that our intuitive attachment to that lower-ranked goal makes us uneasy about that practice. We feel, intuitively, that the goal of meeting needs *itself* somehow suffers in the process of throwing needed resources into the sea. Surely, we say to ourselves, making everyone in society go as short of food as the hungriest member is no way to meet people's needs.

III. CHARACTERIZING REALLY RELATIVE NEEDS

The upshot, I think, is just that people like Townsend have overplayed their hand in pretending that all social needs are relative, in any strong sense. Some are, some are not. And the air of paradox already identified pretty well dissolves once we remind ourselves of that fact.

What is wrong with the strategy of throwing food into the sea in order to meet people's needs is that their need for food is – to a very large extent, anyway – absolute, and not at all relative to how much others in their society are eating. Even if it is impossible for everyone to eat as much as they need it is better – not just better for individuals themselves, but even better in terms of (absolute) needs-satisfaction across the society as a whole – for more people to eat adequately than for fewer to do so.

The question, then, becomes how to separate out needs that are indeed socially relative, in the strong sense. Some, such as Sen, might attempt to do this by distinguishing between needed resources that are of end-use value, on the one hand, and ones that are of merely instrumental value, on the other.¹¹ The point about food – the reason that you are better off with more food than with less, however much or little food others around you might have – is that food is of end-use value to you. You derive benefits from food through the very act of your eating it; you are not depending on any further chain of social consequences, in which food consumption plays some crucial role, in order to get the good in view from the morsel in your mouth.

Thus it would seem that your need for some resource can only be

¹¹ Sen 1983.

socially relative, in the strong sense, if the resource is of merely instrumental value to you. This partitioning of the problem takes us less far than we might hope, however. The trouble is that, as noted in Section I, it is in the nature of needs that needed resources are always of merely instrumental value. In saying that something is needed, you must always say *for what* it is needed.¹² But if that is so, then the distinction between absolute and relative needs cannot be analyzed in terms of a non-existent distinction between needed resources that are of end-use value and ones that are of merely instrumental value. All are instrumental.

Everything will therefore have to depend on the second feature – which was always going to have to be a necessary part of the definition of relative needs, anyway – of competitive utilization. All needed resources are instrumental. But some instruments' usefulness to you is independent of others' stock and use of similar instruments, while some instruments' usefulness is not so independent. Food is an example of the former. It is instrumental in keeping you alive, but your meal's usefulness to you in that regard is independent of anyone else's stock or use of the food in their larder. Money is an example of the latter. It is instrumental, too, but how useful it is to you in pursuing your ends depends on how much others have and use of it in bidding against you for scarce resources that you both desire but cannot both simultaneously enjoy.¹³

Now, the paradox identified above dissolves in the context of needed resources that are instrumental and competitive in their utilization. If your stock and use of resources actually impinges upon others' enjoyment of theirs, then it is easy enough to see how restrictions on your stock and use of resources might enhance overall social enjoyment – what you lose through such restrictions may be more than compensated by what others gain through them. Where instru-

¹² What we call needs may well be those resources that are instrumental for just about anything else one cares to do (Doyal and Gough 1984, p. 14; Daniels 1985, chaps. 1 and 2; Goodin 1988, pp. 35–40); but their instrumental nature is in no way undercut by the universality of the instrument. Means–ends reasoning in general is plagued by the problem that maybe there never is any “end” in the chain. Every end is always, at the same time, merely a means to some higher-order end (Braybrooke and Lindblom 1963).

¹³ Models of this sort include Shubik's (1971) “games of status”; Hirsch's (1976) “positional goods”; and Sen's (1977b; 1981; 1985a; 1985b; Dréze and Sen 1989) “exchange entitlements” as a means of securing “basic capabilities.” Relative needs will be a problem only in circumstances of scarcity. If everyone could have all that they wanted of everything, then there would be no reason for resource disparities to translate into consumption disparities; everyone could, and would, be equally satiated (Ellis and Heath 1983). Still, scarcity in that sense will presumably always be with us.

mental, competitive-utilization resources are concerned, throwing the resources of the rich into the sea may well be a good way to reduce relative social needs overall.

It is worth emphasizing, however, that for the paradox to dissolve in this way the competitiveness must be genuinely objective. It must, somehow, be inherent in the very nature of the goods themselves that one person's utilization of them as instruments toward his or her ends is inherently competitive with another person's utilization of them in similar fashion. Stories that trace the competitiveness to some purely subjective attribute will simply not suffice.

For an example of the latter, inadequate, variation on this theme, consider the familiar story about "relative deprivation." People's felt deprivation is relative to their aspirations and expectations, which are relative in turn to their reference group and average accomplishments within it.¹⁴ A similarly subjective story about relative needs would be that what people think they need is relative to what they expect to be able to do, which is relative in turn to what others around them are doing.

As a story about subjective satisfactions – of either wants or needs – that account is impeccable. Yet surely there is still something paradoxical about any inference that we should therefore destroy needed resources, just to reduce people's subjective anxieties about their relative deprivations. Perhaps the air of paradox persists simply because we do not take subjectivities of this sort altogether seriously. We think that people *are* (objectively) better off with more needed resources, even if others have still more of them, however sensitive the people themselves may be to meaningless relativities. Secondarily, perhaps, we may be particularly tempted to take that hard line on merely subjective deprivation because we suppose that people always have it within their own power to ease any objectively groundless sense of deprivation that they may feel just by revising their aspirations or their choice of who to take as their reference group. Purely subjective problems admit of purely subjective solutions.¹⁵

If the relativities matter not just subjectively but objectively, though, then it is a different story. Instrumental resources that are competitive in their utilization in that way are a wholly legitimate matter of concern, because the more of them that others have the less good mine will objectively do me. Since it is not thinking that made it so in the first place, thinking otherwise will not cease to make it so. It is not within my power to make objective relativities cease to matter, in the way that it is with subjective relativities. With objectively competitive

¹⁴ Hyman 1942. Runciman 1966. Campbell 1972. O. Duncan 1975, p. 273.

¹⁵ Frankfurt 1984, pp. 9–13.

resources, I inevitably am and I must inevitably remain objectively worse off, the relatively more others have of them. From that fact it follows that we may all be made objectively better off – less needy, in ways that objectively matter – by removing resources from the relatively well endowed, even if those resources are then merely thrown into the sea.

IV. POLICY IMPLICATIONS

The argument so far has been designed to establish that, while (*pace* Townsend) not all needs are relative, at least some needs are genuinely relative in the strong sense in which how much you need depends on how much others around you have. Let us call these, generically, needs for “status goods.” The balance of this chapter is devoted to identifying certain specific needs of this sort, and tracing out their policy consequences.

It must be emphasized at the outset, however, that very few goods (hence needs) fall squarely into one category or the other. Most needed resources are valued partly for what they can do for you, independently of others’ utilization of similar resources, and partly for what they can do for you in competition with others. In virtually all the discussions that follow, therefore, I shall persistently be saying “insofar as” the resources are needed as status goods, certain policy consequences follow. There is no presumption that that is the only role that they play, or the only need that they serve. Insofar as they do indeed carry some other value, and serve some other end, then the appropriate social policies *vis-à-vis* those needed resources are to that extent potentially quite different.

With that caveat in place, let me simply catalog what I take to be some of the main headings of social needs, tracing in what ways the needed resources are status goods and what policy prescriptions follow from that fact.

A. Housing

Consider, first, housing needs. It is indisputably true that, to some extent, people’s needs for shelter are absolute, and not at all relative to the housing standards enjoyed by others around them. Of course, what sort of housing you need in the desert is very different from that which you need in the mountains, and what sort you need in damp climates is very different from that which you need in dry ones. But that is just a matter of social relativity in the weaker sense – invariate standards of health and hygiene just generate different implications in different climatic conditions.

There is, then, a "basic need" for housing that is absolute, and not at all socially relative in the strong sense. It is a sad truth, of clear importance for public policy, that not everyone's basic housing needs are presently being met. Even (perhaps especially) in wealthy societies, some people still sleep on sidewalks, and many still live in insanitary environments. Far from all of them do so in any way that could be remotely described as "through their own choosing." The first task for public policy on housing must, unquestionably, be to respond to absolute need of this sort.¹⁶

Beyond that, however, there is a demand (that is often described as a "need") for housing that is indeed socially relative, in the strong sense. Accommodation is often scorned as "substandard," not on the grounds that it is absolutely unsatisfactory (unhealthy, unhygienic, etc.), but merely on the grounds that it falls short of the "standard" typical of that society. Thus, for example, Marx and Engels write:

A house may be large or small: so long as the surrounding houses are equally small, it satisfies all social demands for a dwelling. But let a palace arise beside the little house, and it shrinks from a little house to a hut. . . . However high [the poor man's] house may shoot up in the course of civilization, if the neighbouring palace grows to an equal or even greater extent, the occupant of the relatively small house will feel more and more uncomfortable, dissatisfied and cramped within its four walls.¹⁷

Insofar as housing is indeed a status good of that sort, various policy consequences follow. The first is that relative needs satisfaction will be maximized by a regime that enforces broad equality in housing standards. Over the minimum standard set by people's absolute basic needs for housing, equally small houses are just as good as equally large ones. Within the range here under discussion, it is the equality of everyone's accommodation rather than its absolute size that matters.

The second policy consequence is that, on various other grounds, it is probably socially preferable that housing should be equally modest for all rather than equally grand for all, again once we are over the minimum standard set by absolute basic needs for housing. Contrast, for example, the terraced three-bedroomed houses of an English town with the suburban sprawl of three-bedroomed houses set on a quarter of an acre in Sydney's western suburbs. The joys of living "in the country" are wholly lost when you have to share it with so many

¹⁶ Waldron 1991.

¹⁷ Marx and Engels 1958, vol. 1, pp. 93-4; cf. Streeten 1981, p. 19. There are objective as well as subjective factors at work here. The more space your neighbor's house occupies, the farther you have to walk to visit other neighbors or local shops.

near neighbors; the distances involved make walking to neighborhood shops infeasible and driving, with the attendant traffic congestion at rush hours, mandatory.¹⁶ Thus, insofar as relativities really are all that matter to people in this range of housing choice, everyone would be better off (have their "relative needs" for housing better met) with smaller houses on smaller building plots.

In practical policy terms, this presumably means that relative housing needs would be better met through programs of building restriction rather than by programs of building. Land use controls ought to be used – in the name of meeting relative needs, even – to prevent a socially counterproductive turn in the competition for ever grander houses on ever larger plots in previously unspoiled locations. Negatively, governments ought to refrain from providing infrastructural subsidies (through road building, sewer extensions, etc.) to developers. Positively, governments ought to take steps to protect the "green belts" around urban centers.

None of those policy measures is in the least novel, of course. The novelty lies merely in justifying those measures in the selfsame needs-based terms that are usually employed to justify the house building that I would hope to stop. Depending on the exact nature of the needs, maximizing the supply of dwellings may be counterproductive of satisfaction of housing needs. If the needs in view are socially relative in the strong sense (so the housing is required merely as a status good) then smaller but more equally sized houses will meet that need better than larger and more variably sized ones.

B. Food and clothing

The same that has been said of shelter can be said, *mutatis mutandis*, of the other two members of the classic trio, food and clothing. Both unquestionably, to some extent, address needs that are utterly absolute. Again, how much and what kinds of food you need to eat, and how much and what kinds of clothes you need to wear, depend on vagaries of local climate and so on. But that is social relativism of a relatively superficial sort. At root, one's basic need for certain quantities of food and clothing is surely as absolute as one's need for shelter from the ravages of the environment.

Again, however, people are generally said to have "needs" for food and clothing that go well beyond those absolute necessities. What sort of food you need to eat, and what sort of clothing to wear, is to some extent relative to what others around you are eating and wearing. In

¹⁶ Hirsch 1976, pp. 32-41. K. Jackson 1985.

part, perhaps, that is because of the contribution food and clothing make to one's capacity to utilize one's human capital effectively in the labor market. A well-fed and well-clothed worker can work more efficiently than a less well (albeit perfectly adequately, by absolute standards) fed and clothed worker.¹⁹ In larger part, no doubt, the social relativity of needs for food and clothing derives from the role of those commodities as symbols and signals in every culture.²⁰ Who would trust an ill-clad person who prefers cabbage to caviar with a position of responsibility in our culture?

The upshot, here again, is that food or clothing beyond the absolute basic necessities is largely a "status good," your need for which is a function of how many other people around you already have it. Insofar as that is true, it once again follows that how much of these fancy goods matters less, socially, than how they are distributed. The satisfaction of relative needs, across the whole society, is more a matter of equalizing stocks than of maximizing them.

Social symbolisms are sufficiently flexible, and policy instruments sufficiently blunt, that it is probably hopeless ever to expect to impose even rough equality on consumption of food and clothing. Sumptuary laws have, historically, met with little success since Richard the Lionheart tried to prohibit the extravagant wearing of fur by the Crusaders under his command.²¹ Still, there is one measure that might even nowadays be worth considering.

Highest status typically attaches to imported goods. (That is partly due to their relative scarcity, no doubt; but it is at least partly due to their being exotic, whether or not particularly scarce.) If we want to maximize satisfaction of everyone's relative needs for status goods of food and clothing across the whole society, one way to do it would therefore be to discourage importation of fancy goods, to which high status will quite probably be attached.

The standard way of doing that is to impose high tariffs, of course, on the ground that the more expensive a product is the less of it will be consumed. But that is exactly the wrong way to discourage status goods competition, of course. The higher the price, the more a mark of status the good's consumption is. Thus, instead of taxing French champagne heavily, Australians keen to avoid status competition should ban it altogether. No one will suffer, in absolute terms, from being forced to drink the domestic product instead. There are,

¹⁹ For an elegant formalization see Dasgupta and Ray (1986-7).

²⁰ Douglas and Isherwood 1979.

²¹ Schoeck 1969, p. 261. The requirement, still in force in some places and some schools, that students wear school uniforms similarly avoids status competition based on clothing.

of course, the standard "trade war" rejoinders to this proposal to be considered. My point here is just that, if people's relative needs are what is at issue, that is the best way to meet them.

C. Education

Education falls into the same broad pattern. To a certain extent, more education makes you absolutely better off. You are better able to manipulate nature around you; you enjoy life more, and so on. To some extent, however, one's need for education is socially relative. That is clearly true in the weaker sense, in which how much information you need to have depends on whether you are living in a primitive culture or a technologically sophisticated one. It is also true in the stronger sense, in which how much education you need depends on how much others around you have.

The most interesting way in which education serves as a "status good" has to do with "credential inflation" in the labor market. Suppose that the amount of information and intellectual sophistication required to perform perfectly satisfactorily the job of an entry-level clerk in the civil service is equivalent to three British A-levels, or a high school diploma in the United States. Suppose, however, that employers will always prefer to hire more qualified rather than less qualified candidates for any given job, for any of a variety of reasons.²² Then each person, hoping for a better job, has an overwhelming incentive to acquire qualifications well beyond those that are strictly needed to do the job for which he or she will apply. It is a "prisoner's dilemma" situation among prospective employees, wherein each would be better off if they all refrained, but wherein no one has any incentive so to refrain from acquiring credentials superfluous to the task.²³

Insofar as education is needed only to confer relative advantage in labor market competition (and I hasten to add here, as before, that this is only part of the story), the policy implications are clear, at least for a country with a centralized educational system. We need merely decide how many jobs there are (and are likely to be) requiring what levels of educational attainment; then we ought to restrict entry to those educational schemes to numbers roughly proportionate to the projected need for people with those skills, after allowing for natural

²² There might be good reasons for this practice. Better qualified candidates might perform marginally better or more reliably, or they might be better long-term investments for the employer. More often, though, the practice seems to be employed merely as a tie-breaker among candidates who are by all reasonable criteria equally promising.

²³ Elster 1976a. Hirsch 1976, pp. 41-51.

wastage. In countries leaving such matters to the market, and where people are allowed to buy as much education as they can afford, restricting credential inflation will obviously be harder. But perhaps a first step would be to grade each job, according to the skills strictly required for its performance, and to prohibit employers from discriminating against applicants on the grounds that they do not have educational qualifications in excess of what is deemed necessary.²⁴

The general aim, with education as with other status goods, would be to equalize rather than maximize across society. The rationale, here as with other status goods, is not to promote equality for its own sake but rather as a way of maximizing relative needs-satisfaction across the whole of society. The fewer superfluous credentials each has, the fewer all others will need.

D. Legal aid

On the face of it, "justice" is an absolute virtue. It is better to have more of it than less, for each and every person. Any injustice done to you is a wrong, which is not mitigated by any similar wrongs that might have been done to others around you. From the obvious truth of that proposition, we might (wrongly) infer another. It is always better for people to have more legal services than fewer; more specifically, it is always better for people to have legal representation than not to have it.

Legal representation is commonly described as a social need, and legal aid schemes to provide such representation to those who cannot afford it are generally justified on those grounds. Insofar as it is a matter of providing attorneys for those accused of crimes, such legal aid schemes undoubtedly contribute to "justice" in absolute terms. Where the state will prosecute, it is wrong for the accused to stand undefended. Similarly, perhaps, insofar as it is a matter of defending against a civil action. Where the plaintiff will be represented by counsel, it is a matter of absolute natural justice that the defendant should be likewise.

In other respects, however, legal services are a relative rather than an absolute need. How much legal assistance you need depends on how much others on the other side of the courtroom have to bring to bear. To a large extent, you need lawyers to defend you against lawyers. If no one had lawyers, no one would need lawyers, in civil ac-

²⁴ Perhaps a society so laissez-faire as to leave education to the market would not be tempted to be so interventionist at this point. But the same policy might be justified as a mild form of reverse discrimination, given that education attainment is lowest among minorities (especially racial ones) that have been historically victims of unjustified discrimination.

tions at least. In that sense, legal representation is a "status good," in the technical sense that how much you need depends on how much others around you have.

Insofar as legal aid is indeed a status good of that sort, various more-or-less radical policy prescriptions follow. The rather grander way of putting the point would be to say that we should try to follow the abortive Indian attempt to return (for purposes of civil cases, at least) to less structured forms of pleadings, familiar in England from the early history of Chancery.²² The less grand way of putting the point would be simply to say that we should try to expand upon the model of the "small-claims courts" already in operation in various Anglo-American jurisdictions. The point, in each case, would be that legal representation would then be less necessary, and also less advantageous.

In this case as the others, the general idea is to equalize rather than maximize the supply of legal services across the society. If lawyers are needed only to defend us against other lawyers, then we would be equally (or indeed better) off with fewer rather than more. The less legal aid each has, the less all others will need.

V. CONCLUSION

These brief remarks on a random assortment of policy issues are meant merely to be illustrative. The selection is not systematic; the discussions are not remotely comprehensive on any of the proposals; given how unrealistic, politically, most of the proposals actually are there is little point in fleshing them out further. Still, some more general points of deeper interest emerge from those more particular policy proposals.

There are two components of needs-satisfaction. On the one hand, we have needs that are absolute. The satisfaction that any one person derives from needed resources of this sort is independent of others' utilization of similar resources. On the other hand, we have needs that are relative. There, the satisfaction that any one person derives from needed resources does depend on how many similar resources others have and use. These two components of needs-satisfaction are analytically separable even if in policy terms they are often empirically intertwined.

Insofar as our aim is to maximize satisfaction of absolute needs, the policy prescription is just the familiar one. There, we should indeed maximize people's supplies of needed resources, up to the point that they need no more of them according to standards of absolute need.

²² Goodin 1982b, chap. 4. Galanter 1974.

Insofar as our aim is the satisfaction of relative needs, however, the policy prescriptions are really quite different. There, policy should be guided by two principles. The first is equalization rather than maximization of resources that are needed in this relative way. Maximization of relative needs-satisfaction across the whole society is (given the way "relative needs" in the strong sense have here been defined) equivalent to the minimization of differentials between people's holdings across the whole society. It is the equality of holdings across society, rather than the absolute levels of those holdings, that matters for this purpose. Unlike absolute needs-satisfaction, everyone's being equally poor in these respects is just as good as everyone's being equally rich, so far as relative needs-satisfaction is concerned.

Though those two outcomes are equally good in terms of relative needs-satisfaction, there are other grounds for choosing between them. Insofar as a resource is needed only to procure competitive advantage over others with similar resources, and insofar as competition of that sort is socially counterproductive (eating up resources that could, alternatively, have been put to more productive uses), it is socially preferable that relative needs be satisfied by providing everyone with equally few resources of this sort rather than equally many. Minimizing everyone's equal supply of relatively needed resources, then, is the second principle that should guide policy in this area.

These two principles, taken together, have further implications for the preferred mode of provision of these relatively needed resources. Roughly speaking, the options are to provide them through market mechanisms that have been suitably adjusted or to provide them in kind via direct social provision. Equalization of resources could, in principle, be accomplished through either route. (In practice, of course, there may be far more political resistance to the dramatic equalization of income and assets that would be required to accomplish this via the market mechanism than there would be to the direct provision of nonfungible needed goods in certain more limited categories.) The second goal, however, cannot credibly be pursued through market mechanisms.

It is in the very nature of the competitive market process that each will try to "up the ante" in the status goods sweepstakes, and that each must at least match the other's bid to stay in the competition. This spiraling of the status good competition is socially counterproductive but utterly inevitable, adjust initial market assets as you may. Furthermore, the ordinary market technique for discouraging certain social pursuits will not work with status goods. Whereas the higher the price of most goods the fewer of them are consumed, with status goods the higher their price the more a mark of higher status (and hence the more valuable a token in social competition) they are. Thus,

equalizing relatively needed resources at minimal levels is impossible through market mechanisms.

There are many things that can be said in favor of a policy of meeting social needs through direct social provision of the needed commodities.⁴⁸ Here, I have added yet one more. Where the needs are relative, in the strong sense, the great advantage of the state over the market as provider of those resources is that the state can (at least in principle) say "no" in a way that the market simply cannot. With relatively needed resources – ones needed only to procure competitive advantage, in a competition that is itself socially counterproductive – uniform provision, at lower levels, will achieve all the good without doing any of the harm that would be done by market-style maximization of those resources. The real advantage of direct state provision for those sorts of needs is that the state can, in a way the market cannot, actually restrict supply of those sorts of resources.

Whether there will actually be the political will for the state to say "no" in this way is, perhaps, an open question. There should be. If the competition in status goods is genuinely counterproductive – if everyone really would be better off if everyone (themselves included) were prevented from pursuing them – then there should be a substantial majority, approaching unanimity, in favor of banning the counterproductive competition. But that assumes that everyone is realistic in assessing this as a competition that no one can win. Psychometric studies militate against any such easy assumptions. It is a common phenomenon for people to think that they are better than average, or in the top half of the distribution, on any range of favored attributes you care to mention. And if that tendency is so strong as to encourage large numbers of people to think that, against all the odds, they are likely to be among the very few winners of this status-goods competition, then the natural, rational majority for banning such competitions might start to slip away.

Of course, needed resources are never wholly of one sort or wholly of the other – neither wholly absolute nor wholly relative. So these conclusions must be phrased with care. What my argument implies is that, apropos of any needed resource that has a "status good" relativistic component (and most seem to have such a component, at least across part of their ranges), there is something to be said for uniformly minimal supply achieved through direct state provision; and there is

⁴⁸ Politically, that practice plays on the fact that we tend to be "specific egalitarians" and to care about equality in the distribution of certain commodities more than others or than money in general (Tobin 1970; Walzer 1983). Economically, in-kind transfers are more efficient in satisfying demand where there is greater variation in people's income than in their tastes for the commodities being distributed (Browning 1975; Weitzman 1977).

something, conversely, to be said against markets which, adjust them as we may, will always tend to maximize supply of those resources. That there is "something to be said" for such a policy does not imply that that is a conclusive consideration. Countervailing considerations must always be borne in mind before final policy recommendations can be given. Still, that is an important, if partial, finding.

It is important, if for no other reason, because it flies so squarely in the face of conventional wisdom. Ordinarily we think that maximization of needed resources is socially desirable – as indeed it is, where absolute needs are concerned. Writers employing a relative standard of needs help themselves willy-nilly to that standard presumption. They tend to assume, without argument, that the right response to relative deprivation is to try to bring everyone up to the standards enjoyed by the median member of that society. Where the needs at issue are genuinely relative, though, that is not necessarily – and, if I am right, is necessarily not – the right response. Minimization rather than maximization of those resources might be socially optimal. There, less really may be more, and there we may well want to rely on the state to enforce those minimalist strictures upon ourselves.²⁷

²⁷ I am grateful for comments and criticisms from participants in the Paris ECPR workshop, especially Keith Dowding, Heiner Ganßmann, Philippe Van Parijs and Alan Ware.

International ethics

What is so special about our fellow countrymen?

There are some "general duties" that we have toward other people, merely because they are people. Over and above those, there are also some "special duties" that we have toward particular individuals because they stand in some special relation to us. Among those are standardly supposed to be special duties toward our families, our friends, our pupils, our patients. Also among them are standardly supposed to be special duties toward our fellow countrymen.

When reflecting upon what "special treatment" is due to those who stand in any of those special relations to us, ordinarily we imagine that to be especially *good* treatment. Close inspection of the case of compatriots reveals that this is not completely true, however. At least in some respects, we are obliged to be more scrupulous – not less – in our treatment of nonnationals than we are in our treatment of our compatriots.¹

Of course, it is a politically important result in itself to show that some of our general duties to those beyond our borders are sometimes more compelling, morally speaking, than some of our special duties to our fellow citizens. More to the point of the present book, this finding has the further effect of forcing us to reconsider the bases of our special duties to compatriots – with yet further political consequences. Morally, what ultimately matters is not nationality *per se*. It is instead some other feature that is only contingently and imperfectly associated with shared nationality. This feature may sometimes be found among foreigners as well. When it is, we would have duties toward those foreigners that are similar in their form, their basis and

¹ I shall be making no distinction in this chapter between "state" and "nation," between "citizenship" and "nationality." Those terms will here be used interchangeably.

perhaps even their strength to the duties that we ordinarily acknowledge toward our fellow countrymen.

I. THE PARTICULARIST'S CHALLENGE

A

Modern moral philosophy has long been insistently universalistic. Of course universal laws play themselves out in different ways in different venues and demand different things from differently placed agents. But while their particular applications might vary, the ultimate moral principles, their form and content, has long been regarded as essentially invariant across people. The same basic precepts apply to everyone, everywhere, the same.

A corollary of this universality is impartiality.¹ It has long been supposed that moral principles – and therefore moral agents – must, at root, treat everyone the same. Of course, here again, basic principles that are perfectly impartial can (indeed, usually will) play themselves out in particular applications in such a way as to allow (or even to require) us to treat different people differently. But the ultimate principles of morality must not themselves play favorites.

On this much, at least, utilitarians and Kantians – the great contending tribes of modern moral philosophy – can agree. Everyone counts for one, no one for more than one, in the Benthamite calculus. While as an upshot of those calculations some people might gain and others lose, the calculations themselves are perfectly impartial. So too with Kant's Categorical Imperative. Treating people as ends in themselves, and respecting the rationality embodied in others, may require us to do different things to, for, or with different people. But that is not a manifestation of any partiality between different people or their various projects. It is, instead, a manifestation of our impartial respect for each and every one of them.

Furthermore, this respect for universality and impartiality is no mere quirk of currently fashionable moral doctrines. Arguably, at least, those are defining features of morality itself. That is to say, they arguably must be embodied in any moral code in order for it to count as a moral code at all.

¹Or so it is standardly supposed. Actually there could be a "rule of universal partiality" (e.g., "everyone ought to pursue his own interests" or "everyone ought to take care of his own children"). A variant of this rule figures largely in my argument in Section V.

B

Despite this strong attachment to canons of universality and impartiality, we all nonetheless ordinarily acknowledge various special duties. These are different in content and form from the general duties that universalistic, impartial moralities would most obviously generate for us. Whereas our general duties tell us how we should treat anyone, and are hence the same toward everyone, special duties vary from person to person. In contrast to the universality of the general moral law, some people have special duties that other people do not. In contrast to the impartiality of the general moral law, we all have special duties to some people that we do not have to others.¹

Special duties, in short, bind particular people to particular other people. How this particularism of special duties fits with the universality and impartiality of the general moral law is problematical. Some say that it points to a whole other branch of the moral law, not captured by any of the standard canons. Others, Kantians and utilitarians among them, say that it is derivative in some way or another from more general moral laws. Yet others say that this particularism marks the limits of our psychological capacities for living up to the harsh standards that the general moral law sets for us.⁴

Be all these foundational questions as they may, it is not hard to find intuitively compelling examples of special duties that we would all acknowledge. At the level of preposterous examples so favored among philosophers, consider this case. Suppose your house is on fire. Suppose two people are trapped in the fire, and you will clearly have time to rescue only one before the roof collapses killing the other. One of those trapped is a great public benefactor who was visiting you. The other is your own mother. Which should you rescue?

This is a story told originally by an impartialist, William Godwin. Being a particularly blunt proto-utilitarian, he had no trouble plunking for the impartialist position: "What magic is there in the pronoun 'my' that should justify us in overturning the decisions of impartial truth?"² Nowadays, however, it is a story told more often against impartialists. Few, then or now, have found themselves able to accept

¹ The terms and surrounding analysis of "special" and "general duties" derive from Hart (1955).

² See Goodin 1985c, chap. 1, and references therein. The strongest arguments for such partiality, focusing on the need to center one's sense of self, through personal attachments to particular people and projects (B. Williams 1981, chap. 1), apply most strongly to more personal links. They apply only very weakly if at all to impersonal links through shared race or nationality (Cottingham 1986, pp. 370-1).

³ Godwin [1793] 1971, bk. 2, chap. 2.

the impartialist conclusion with quite such equanimity as Godwin. Many regard the example as a *reductio ad absurdum* of the impartialist position. And even those who want to stick up for the impartialist side are obliged to concede that impartialists have a case to answer here.⁶

But the debate is not confined to crazy cases like that one. In real life, just as surely as in moral fantasies, we find ourselves involved in special relations of all sorts with other people. And just as we intuitively feel that we should save our own mothers rather than Archbishop Fenelon in Godwin's example, so too do we intuitively feel we should show favoritism of some sort to all those other people likewise. The "mere enumeration" of people linked to us in this way is relatively uncontentious and has changed little from Sidgwick's day to Parfit's. Included in both their lists are family, friends, benefactors, clients, and co-workers and – especially important, in the present context – compatriots.⁷

Intuitively, we suppose that, on account of those special relations between us, we owe all of those people special treatment of some sort or another: special "kindnesses," "services" or "sacrifices"; "we believe that we ought to try to give them certain kinds of benefit." According to Parfit, "Common-Sense Morality largely consists in such obligations"; and, within commonsense morality, those obligations are particularly strong ones, capable of overriding (at least at the margins) our general duties to aid strangers.⁸

C

Here, I do not propose to focus (initially, at least) upon the precise strength of those duties. Rather, I want to direct attention to their general tendency. Notice that there is a presumption, running through all those standard discussions of special duties, that the special treatment due to those who are linked to us by some special relation is especially *good* treatment. We are said to be obliged to do more for those people than for unrelated others in an effort to spare them harm or to bring them benefits. To those who stand in some special relation to us, we are said to owe special "kindnesses," "services" or "sacrifices."

That assumption seems to me unwarranted. Agreed, special relations do sometimes permit (and sometimes even require) us to treat those specially related to us better than we need to, absent such a link.

⁶ See, respectively, B. Williams 1981, pp. 17–18, and Hare 1981, p. 138.

⁷ Sidgwick [1874] 1907, bk. 3, chap. 4, sec. 3. Parfit 1984, pp. 95, 485.

⁸ Sidgwick [1874] 1907, bk. 3, chap. 4, sec. 3. Parfit 1984, pp. 95, 485.

Other times, however, special relations permit (and perhaps even sometimes require) us to treat those thus linked to us worse than we would be obliged to treat them, absent such a link. Exploring how that is so, and why, sheds light upon the true nature and strength of special duties. It also, not incidentally, limits the claims for exclusive special treatment that can be entertained under that heading.

II. THE CASE OF COMPATRIOTS

When discussing what special claims compatriots, in particular, have against us, it is ordinarily assumed that we owe more to our fellow countrymen and less to foreigners. The standard presumption is that "compatriots take priority" over foreigners, "at least in the case of duties to aid"; "the state in determining what use shall be made of its own moneys, may legitimately consult the welfare of its own citizens rather than that of aliens."⁹ Thus, it makes a salutary start to my analysis to recall that, at least with respect to certain sorts of duties, we must be more scrupulous – not less – in our treatment of foreigners.

In the discussion that follows, "we" will be understood to mean "our community, through its sovereign representatives." In discussing what "we" may and may not do to people, I shall require some rough-and-ready guide to what our settled moral principles actually are. For these purposes, I shall have recourse to established principles of our legal codes. Although the correspondence is obviously less than perfect, presumably the latter at least constitute a rough approximation to the former. Public international law will be taken as indicative of what we may do to foreigners, domestic public law as indicative of what we may do to our compatriots. In both cases, the emphasis will be upon customary higher law rather than upon merely stipulative codes (treaties, statutes, etc.).¹⁰

Consider, then, all these ways in which we must treat foreigners in general better than we need to treat our compatriots:¹¹

⁹ Quoting Shue (1980, p. 132) and Cardozo (1915; p. 164/437), respectively. This report of the conventional wisdom is echoed by Nagel (1978, p. 81), Beitz (1979, p. 163) and Goodin (1985c, chaps. 1 and 2), with only Cardozo among them being uncritically accepting of it.

¹⁰ Unlike stipulative law, which might be made by a small body of people on the spur of the moment, customary law represents the settled judgments of a great many people over some long period. It is thus a better guide to our collective settled judgments on these issues (Goodin 1976, chap. 7; 1985c, chap. 5).

¹¹ Specific contractual or treaty commitments may stipulate better treatment, or permit worse, or both in different respects. The principles set out in the text constitute the normative background against which such contracts or treaties are negotiated, however.

Example a. We, through our public officials, may quite properly take the property of our fellow citizens for public purposes, provided they are duly compensated for their losses; this is especially true if the property is within our national boundaries but is even true if it is outside them. We cannot, however, thus commandeer an identical piece of property from a foreigner for an identical purpose in return for identical compensation. This is especially true if the property is beyond our borders; but it is even true if the property is actually in our country, in transit.¹²

Example b. We can conscript fellow citizens for service in our armed forces, even if they are resident abroad.¹³ We cannot so conscript foreign nationals, even if they are resident within our own country.¹⁴

Example c. We can tax fellow citizens, even if they are resident abroad. We cannot so tax foreigners residing abroad on income earned abroad.¹⁵

Example d. We can dam or divert the flow of a river lying wholly within our national territory to the disadvantage of fellow citizens living downstream. We may not so dam or divert rivers flowing across international boundaries to the disadvantage of foreigners downstream.¹⁶

Example e. We can allow the emission of noxious factory fumes that damage the persons or property of fellow citizens. We may

¹² Fisher 1965, sec. 185c. This "right of safe passage," for people and goods in transit for purposes of commerce or study, dates to Grotius ([1625] 1925, bk. 2, chap. 2, secs. 13-15), Wolff ([1749] 1934, sec. 346) and Vattel ([1758] 1863, bk. 2, chap. 10, sec. 132). The problem in expropriating foreigners' property outside our borders does not derive from problems in appropriating a piece of another nation's territory. The same problem remains with a piece of movable property (such as a ship - the British government may requisition a ship of British registry even if it is lying in Dutch waters, but may not except in dire emergency requisition a Dutch ship even if it is lying in British waters).

¹³ Oppenheim 1955, vol. 1, p. 288. This and the similar result in example c follow from the fact that the state enjoys continuing "personal" sovereignty over aliens within its borders (Suárez 1612, chap. 30, sec. 12).

¹⁴ Oppenheim 1955, chap. 1, p. 288. The United States is exceptional in conscripting aliens who are permanently resident into its armed forces (Bickel 1975, p. 49). The long-standing rule in international law is that, while resident aliens may be required to help with police, fire and flood protection, foreigners are exempt from serving in the militia (Vattel [1758] 1863, bk. 2, chap. 8, secs. 105-6).

¹⁵ Oppenheim 1955, vol. 1, p. 288. Bickel 1975, p. 48. Earlier writers argued that even resident aliens should be exempt from certain sorts of taxes (Wolff [1749] 1934, sec. 324; Vattel [1758] 1863, bk. 2, chap. 8, sec. 106).

¹⁶ Oppenheim 1955, vol. 1, pp. 290-1, 348, 475.

not do so if those fumes cross international frontiers, causing similar damage to the persons or property of foreigners there.¹⁷

Example f. We may set arbitrarily low limits on the legal liability of manufacturers for damages done by their production processes or products domestically to our fellow citizens. We may not so limit the damage recoverable from them for harm done across international boundaries to foreigners.¹⁸

Example g. According to international law, we may treat our fellow citizens "arbitrarily according to [our own] discretion." To aliens within our national territory, however, we must afford their persons and property protection "in accordance with certain rules and principles of international law," that is, "in accordance with ordinary standards of civilization."¹⁹ Commentators on international law pointedly add, "It is no excuse that [a] State does not provide any protection whatever for its own subjects" in those respects.²⁰

These are all examples of ways in which we must treat foreigners better than compatriots. In a great many other respects, of course, the conventional wisdom is perfectly right that we owe better treatment to our compatriots than we do to foreigners. For example, we have a duty to protect the persons and property of compatriots against attack,

¹⁷ Oppenheim 1955, vol. 1, p. 291.

¹⁸ "It is," says Oppenheim (1955, vol. 1, p. 350), "a well-established principle that a State cannot invoke its municipal legislation as a reason for avoiding its international obligations."

¹⁹ Oppenheim 1955, vol. 1, pp. 686-7; see further vol. 1, pp. 190, 350, 641. Brierty 1936, pp. 172ff. Even U.S. black-letter law holds that "conduct attributable to a foreign state and causing injury to an alien is wrongful under international law . . . if it departs from the international standard of justice" (Fisher 1965, sec. 165 [1](a)).

²⁰ Oppenheim 1955, vol. 1, pp. 687-8. Elsewhere, Oppenheim (1955, vol. 1, p. 641) draws attention to the "paradoxical result" that "individuals, when residing as aliens in a foreign state, enjoy a measure of protection . . . denied to nationals of a State within its own territory." From the early modern period through the 1940s, many international lawyers held that since the state could refuse admission to aliens altogether, it could impose any conditions it liked upon their remaining in the country, however discriminatory and however uncivilized that treatment may be; at the very least, aliens they would say are not wronged if treated no worse than nationals (Wolff [1749] 1934, sec. 302; Vattel [1758] 1863, bk. 2, chap. 8; Sidgwick 1891, pp. 235-6; Brownlie 1966, p. 425; cf. Suárez 1612, chap. 33, sec. 7). By now, however, it is decidedly the "prevailing rule" in international law that "there is an international standard of justice that a state must observe in the treatment of aliens, even if the state does not observe it in the treatment of its own nationals, and even if the standard is inconsistent with its own law" (Fisher 1965, sec. 165, comment a; Sohn and Baxter 1961, pp. 547-8).

even when they are abroad.¹¹ Absent treaty obligations, we have no such duty to protect noncitizens beyond our borders. We have a duty – morally, and perhaps even legally – to provide a minimum level of basic necessities for compatriots. Absent treaty obligations, we have no such duty – legally, anyway – to assist needy noncitizens beyond our borders.

Even within our borders, we may treat citizens better in all sorts of ways than we treat noncitizens, just so long as some “reasonable” grounds for those discriminations can be produced and just so long as the protection we provide aliens’ persons and property comes up to minimal internationally acceptable standards.¹² Not only are aliens standardly denied political rights, like voting and office-holding, but they are also standardly excluded from “public service.” This has, in the past, been interpreted very broadly indeed. In the United States, an alien could have been debarred from being an “optometrist, dentist, doctor, nurse, architect, teacher, lawyer, policeman, engineer, corporate officer, real estate broker, public accountant, mortician, physiotherapist, pharmacist, pedlar, pool or gambling-hall operator’;”¹³ in the United Kingdom the range of prohibited occupations has included harbor pilots, masters of merchant ships, and skippers of fishing vessels.¹⁴ Besides all those quasi-public functions from which aliens are excluded, they also suffer other disadvantages of a purely material sort. Perhaps the most significant among them are the rules found in some states denying aliens the right to own land.¹⁵ All this can be perfectly permissible, both under international law and under higher domestic law.

Thus, the situation is very much a mixed one. Sometimes we are indeed permitted (sometimes even required) to treat our fellow citizens better than we treat those who do not share that status with us. Other times, however, we are required to treat noncitizens better than we need to treat our own fellow citizens.

I pass no judgment on which pattern, on balance, predominates. The point I want to make here is merely that the situation is much more mixed than ordinary philosophical thinking on special duties leads us to expect. That in itself is significant, as I shall now proceed to show.

¹¹ These however are obligations of customary and higher domestic law, unenforceable under international law (Oppenheim 1955, vol. 1, pp. 686–7).

¹² Suárez 1612, chap. 33, sec. 7. Wolff [1749] 1934, sec. 303. Sidgwick 1891, p. 235. Brierly 1936, pp. 172–3. Oppenheim 1955, vol. 1, pp. 689–91. Brownlie 1966, pp. 424–48. Rosenberg 1977. Corwin 1978, pp. 90–2; 1980, pp. 159–61. Anonymous 1983.

¹³ Bickel 1975, pp. 45–6. See also Corwin 1978, pp. 90–2; Chase and Ducat 1980, pp. 159–61; and anonymous 1983.

¹⁴ Brierly 1936, p. 173. Oppenheim 1955, vol. 1, p. 690.

¹⁵ Brierly 1936, p. 173. Bickel 1975, p. 46. Anonymous 1983, pp. 1300–1.

III. SPECIAL DUTIES AS MAGNIFIERS AND MULTIPLIERS

In attempting to construe the effect that special relationships have on our moral duties, commonsense morality tends to employ either of two basic models (or both of them – they are nowise incompatible). On the face of things, these two models can only offer reinforcing interpretations for the same one half of the phenomenon observed in Section II. Digging deeper to see how such models might account for that other half of the phenomenon drives us toward a model that is even more deeply and familiarly flawed.

A

One standard way of construing the effect of special relationships on our moral duties is to say that special relationships “merely magnify” preexisting moral duties. That is to say, they merely make more stringent duties which we have, in weaker form, *vis-à-vis* everyone at large; or, “imperfect duties” are transformed by special relationships into “perfect” ones. Thus, perhaps it is wrong to let anyone starve, but it is especially wrong to let kin or compatriots starve. And so on.

That kind of account fits only half the facts, as sketched in Section II, though. If special relationships were merely magnifiers of preexisting duties, then the magnification should be symmetrical in both positive and negative directions. Positive duties (i.e., duties to provide positive assistance to others) should become more strongly positive *vis-à-vis* those linked to us by some special relationship. Negative duties (i.e., duties not to harm others) should become more strongly negative *vis-à-vis* those linked to us by some special relationship. When it comes to our duties in relation to compatriots, however, the former is broadly speaking true, while the latter is not.

It is perfectly true that there is a variety of goods that we may or must provide to compatriots that we may at the same time legitimately deny to nonnationals (especially nonresident nonnationals). Rights to vote, to hold property, and to the protection of their persons and property abroad are among them. In the positive dimension, then, the “magnifier” model is broadly appropriate.²⁸

In the negative dimension, it is not. All the examples *a* through *f* in Section II point to ways in which we may legitimately impose bur-

²⁸ Example *g* above arguably does not fit this pattern. It depends on whether we construe this as a positive duty to provide aliens with something good (“due process of law”) or as a negative duty not to do something bad to them (“deny them due process of law”); this in turn depends on where we set the baseline of how well off they would have been absent our intervention in the first place.

dens upon compatriots that may not properly be imposed upon non-nationals (especially nonresident nonnationals). We may poison our compatriots' air, stop their flow of water, deprive them of liberty by conscription, deny them legal remedies for damage to their persons and their property – all in a way that we cannot do to nonresident nonnationals. If anything, it is our negative duties toward nonnationals, not our negative duties toward compatriots, that are here magnified.

B

A second way of construing the effect of special relationships on our moral duties is to say that special relationships “multiply” as well as magnify preexisting duties. That is to say, special relationships do not just make our ordinary general duties particularly stringent in relation to those bound to us by some special relationship; they also create new special duties, over and above the more general ones that we ordinarily owe to anyone and everyone in the world at large. Thus, contracts, for example, create duties *de novo*. I am under no general duty, strong or weak, to let Dick Merelman inhabit a room in my house; that duty arises only when, and only because, we sign a lease. The special (here, contractual) relationship has created a new duty from scratch.

The “multiplier” model bolsters the “mere magnifier” model’s already broadly adequate account of why we have especially strong positive duties toward those linked to us by some special relationship. Sometimes those special relationships strengthen positive duties we owe, less strongly, to everyone at large. Other times, special relationships create new positive duties that we owe peculiarly to those linked to us. Either way, we have more and stronger positive duties toward those who stand in special relationships to us than we do to the world at large. And that broadly fits the pattern of our special duties vis-à-vis compatriots, as revealed in Section II.

On the face of it, though, it is hard to see how this multiplier model can account for the weakening of negative duties toward compatriots observed there. If special relationships multiply duties, then we would ordinarily expect that that multiplication would produce more new duties in each direction. Consider the paradigm case of contracts. Sometimes contracts create new special duties enjoining us to help others in ways that we would not otherwise be bound to do. Other times, contracts create new special duties enjoining us not to harm others (e.g., by withdrawing trade, labor, or raw materials) in ways that we would otherwise be at liberty to do. It is hard, on the face of it, at least, to see what the attraction of special duties would be – either

for agents who are anxious to incur them or for philosophers who are anxious to impose them – if they make people worse off, opening them up to new harms from which they would otherwise be protected.

Yet, judging from examples *a* through *f* in Section II, that is precisely what happens in the special relationship between compatriots. Far from simply creating new negative duties among compatriots, that special relationship seems sometimes to have the effect of canceling (or at least weakening or mitigating) some of the negative duties that people owe to others in general. That hardly looks like the result of an act of multiplication. Ordinarily, we would expect that multiplication should produce more – not fewer – duties.

C

Digging deeper, we find that there may be a way to explain why special relationships have this curious tendency to strengthen positive duties while weakening negative ones. This model quickly collapses into another, more familiar one – and ultimately falls prey to the same objections standardly lodged against it, as Section IV will show. Still, it is worth noting how quickly all the standard theories about special duties, when confronted with certain elementary facts about the case of compatriots, collapse into that familiar and flawed model that ordinarily we might have regarded as only one among many possible ways of filling out those theories.

The crucial move in reconciling standard theories about special duties with the elementary facts about compatriots laid out in Section II is just this – whether special relationships multiply duties or merely magnify them, the point remains that a relationship is inherently a two-way affair. The same special relation that binds me to you also binds you to me. Special duties for each of us will usually follow from that fact.⁷⁷

Each of us will ordinarily benefit from others' being bound by those extra (or extra strong) duties to do for us things that they are not obliged (or not so powerfully obliged) to do for the world at large. Hence the apparent "strengthening" of positive duties in consequence of special relationships.

Each of us will also ordinarily suffer from those extra (or extra strong) duties imposing an extra burden on us. Hence the apparent "weakening" of negative duties in consequence of the special rela-

⁷⁷ There are of course some unilateral power relations – like that of doctor and patient, or parent and child – that might imply special duties for one but not the other party to the relationship (Goodin 1985c).

tionship. We may legitimately impose burdens on those standing in special relationships to us that we may not impose on those in no special relation to us, merely because we have special rights against them, and they have special duties toward us. Those extra burdens upon them are no more, and no less, than the fair price of our being under special duties to provide them with valued assistance.

Many of the findings of Section II lend themselves quite naturally to some such interpretation. When we say that compatriots may have their incomes taxed, their trucks commandeered, or other liberties curtailed by conscription, that is surely to say little more than that people may be required to do what is required in order to meet their special duties toward their fellow citizens – duties born of their fellow citizens' similar sacrifices to benefit them. When we say that nonnationals (especially nonresident nonnationals) may not be treated in such ways, that is merely to say that we have no such special claims against them nor they any such special duties toward us.

Others of the examples in Section II (especially examples *d* through *g*) do not lend themselves quite so obviously to this sort of analysis. But perhaps, with a sufficiently long story that is sufficiently rich in lurid details, we might be persuaded that polluting the air, damming rivers, limiting liability for damages, and denying people due process of law really is to the good of all; and suffering occasional misfortunes of those sorts really is just the fair price that compatriots should be required to pay for the benefits that they derive from those broader practices.

Notice that, given this account, the motivational quandary in Section IIIB disappears. People welcome special relationships – along with the attendant special rights and special duties (i.e., along with the strengthening of positive duties and the weakening of negative ones) – because the two come as part of an inseparable package, and people are on net better off as a result of it. That is just to say, their gains from having others' positive duties toward them strengthened exceeds their costs from having others' negative duties toward them weakened, and it is impossible for them to realize the gains without incurring the costs.

Notice, however, how quickly these standard theories of how special relationships work on our moral duties – the magnifier and the multiplier models – have been reduced to a very particular theory about “mutual-benefit societies.” Initially, the magnifier and multiplier theories seemed to be much broader than that, open to a much wider variety of interpretations and not committing us to any particular theory about why or how the “magnification” or “multiplication” of duties occurred. Yet if those models are to fit the elementary facts about duties toward compatriots in Section II at all, they must fall

back on a sort of mutual-benefit logic that provides a very particular answer to the question of how and why the magnification or multiplication of duties occurred. As Section IV will show, that is not an altogether happy result.

IV. THE MUTUAL-BENEFIT-SOCIETY MODEL

According to the conventional wisdom about international relations, we have a peculiarly strong obligation to leave foreigners as we found them. "Nonintervention" has long bid fair to constitute the master norm of international law.²⁸ That is not to say that it is actually wrong to help foreigners, of course. It is, however, to say that it is much, much more important not to harm them than it is to help them. Where compatriots are concerned, almost the opposite is true. According to the flip side of that conventional wisdom, it is deeply wrong to be utterly indifferent toward your fellow countrymen; yet it is perfectly permissible for fellow countrymen to impose hardships on themselves and on one another to promote the well-being of their shared community.

Perhaps the best way to make sense of all this is to say that, within the conventional wisdom about international relations, nation-states are conceptualized as ongoing mutual-benefit societies. Within mutual-benefit-society logic, it would be perfectly permissible to impose sacrifices on some people now so that they themselves might benefit in the future; it may even be permissible to impose sacrifices on some now so that others will benefit, either now or in the future.

Precisely what sorts of contractarian or utilitarian theories are required to underpin this logic can be safely left to one side here. It is the broad outline, rather than the finer detail, that matters for present purposes. The bottom line is always that, in a mutual-benefit society, imposing harms is always permissible – but only on condition that some positive good comes of it, and only on condition that those suffering the harm are in some sense party to the society in question.

Suppose, now, that national boundaries are thought to circumscribe mutual-benefit societies of this sort.²⁹ Then the broad pattern of duties

²⁸ Medieval prescriptions along these lines were strengthened in turn by each of the early modern international lawyers – Grotius, Wolff and Vattel – so that by the time of Sidgwick (1891, p. 231) it could be said without equivocation that the "principle of mutual non-interference" is "the fundamental principle" of international morality. It remains so to this day in the view of most lawyers and many philosophers (Walzer 1977; 1980).

²⁹ This follows from Hume's (1739, bk. 3, pt. 2, sec. 2; 1777, sec. 3, pt. 1) account of the "circumstances of justice." Many modern writers follow him in this analysis (Rawls 1971, pp. 126–30), especially in its application to international affairs (Beitz 1979, pp. 143–53; 1983, p. 595; Honoré 1982, p. 154; cf. Wolff [1749] 1934).

toward compatriots and foreigners, respectively, as described in Section II, becomes perfectly comprehensible. In dealing with other people in general (i.e., those who are not party to the society), the prime directive is "avoid harm." Those outside our mutual-benefit society ought not be made to bear any of our burdens; but neither, of course, have they any claim on any of the benefits which we have produced for ourselves, through our own sacrifices. In dealing with others in the club (i.e., compatriots), positive duties wax while negative ones wane. It is perfectly permissible to impose hardships, so long as some positive good somehow comes of doing so; but the point of a mutual-benefit society, in the final analysis, must always be to produce positive benefits for those who are party to it.

There are many familiar problems involved in modeling political communities as mutual-benefit societies.¹⁰ The one to which I wish to draw particular attention here is the problem of determining who is inside the club and who is outside it. Analysis of this problem, in turn, forces us back to the foundational questions skirted at the outset of the article. These will be readdressed in Section V, where I construct an alternative model of special duties as not very special, after all.

From the legalist perspective that dominates discussion of such duties, formal status is what matters. Who is a citizen? Who is not? That, almost exclusively, determines what we may or must do to people, *qua* members of the club.

Yet formal status is only imperfectly and contingently related to who is actually generating and receiving the benefits of the mutual-benefit society. The mismatch is most glaring as regards resident aliens. They are often net contributors to the society, yet they are equally often denied its full benefits.¹¹ The mismatch also appears only slightly less glaringly, as regards natural-born citizens who retain that status although they are and will inevitably (because, e.g., severely handicapped) continue to be net drains on the mutual-benefit society.¹²

In its starkest form, mutual-benefit-society logic should require that people's benefits from the society be strictly proportional to the contributions they have made toward the production of those benefits. Or, minimally, it should require that no one draw out more than he has paid in. The allocation of any surplus created by people's joint

¹⁰ These are illustrated in their particular application to international obligations by Barry (1982, pp. 225-43) and Goodin (1985c, pp. 154-60).

¹¹ Both domestic and international law go some way toward recognizing that in many respects resident aliens are more like citizens than they are like nonresident aliens. But by and large that recognition results not so much in their enjoying the same benefits as in their being forced to bear the same burdens as citizens (Oppenheim 1955, vol. 1, pp. 680-1).

¹² Barry 1979b, pp. 68-9. Goodin 1982b, pp. 77-9.

efforts may be left open. On that logic, we have special duties toward those whose cooperation benefits us, and to them alone. That they share the same color passport – or, indeed, the same parentage – is related only contingently, at best, to that crucial consideration.

It may well be that mutual-benefit logic, in so stark a form, is utterly inoperable. Constantly changing circumstances mean that everything from social insurance to speculative business ventures might benefit us all in the long run, even if at any given moment some of them constitute net drains on the system. And lines on the map, though inherently arbitrary at the margins, may be as good a way as any of identifying cheaply the members of a beneficially interacting community. So we may end up embracing the formalistic devices for identifying members of the mutual-benefit society, knowing that they are imperfect second-bests but also knowing that doing better is impossible or prohibitively expensive.

The point remains, however, that there are some clear, straightforward adjustments that ought to be made to such “first stab” definitions of membership, if mutual-benefit logic underlay membership. That they are not made – and that we think at least one of them ought not be made – clearly indicates that it is not mutual-benefit logic that underlies membership, after all.

Reflect, again, upon the case of resident aliens who are performing socially useful functions over a long period of time. Many societies egregiously exploit “guest workers,” denying them many of the rights and privileges accorded to citizens despite the fact that they make major and continuing contributions to the society. Politically and economically, it is no mystery why they are deprived of the full fruits of their labors in this way.¹³ But if the moral justification of society is to be traced to mutual-benefit logic, that is transparently wrong. The entry ticket to a mutual-benefit society should, logically, just be conferring net benefits on the society.¹⁴ That membership is nonetheless denied to those who confer benefits on the society demonstrates that the society is not acting consistently on that moral premise. Either it is acting on some other moral premise or else it is acting on none at all (or none consistently, which morally amounts to the same).

Or consider, again, the case of the congenitally handicapped. Though born of native parents in the homeland, and by formalistic criteria therefore clearly qualified for citizenship, such persons will

¹³ The argument here parallels that for supposing that, if a workers' cooperative needs more labor, it would hire workers rather than selling more people shares in the cooperative (Meade 1972; Miller 1981).

¹⁴ The “participation” model of citizenship, popular among certain international lawyers (anonymous 1983, pp. 1303–11; Schuck 1984) is obviously a close cousin to this mutual-benefit-society model.

never be net contributors to the mutual-benefit society. If it were merely the logic of mutual benefit that determined membership such persons would clearly be excluded from the benefits of society. (If their parents cared about them, they could give them some of *their* well-earned benefits.) Yet that does not happen, no matter how sure we are that handicapped persons will be net drains on the society for the duration of their lives. And most of us intuitively imagine that it is a good thing, morally, that it does not happen. Thus, society here again seems to be operating on something other than mutual-benefit logic; and here, at least, we are glad that it is.

V. THE ASSIGNED RESPONSIBILITY MODEL

The magnifier, multiplier and mutual-benefit-society models all take the specialness of special duties particularly seriously. They treat such duties as if they were, at least in (large) part, possessed of an independent existence or of an independent moral force. I want to deny both of those propositions.

My preferred approach to special duties is to regard them as being merely "distributed general duties." That is to say, special duties are in my view merely devices whereby the moral community's general duties get assigned to particular agents. For this reason, I call mine an "assigned responsibility" model.

This approach treats special duties as much more nearly derivative from general duties than any of the other approaches so far considered. Certainly it is true that, on this account, special duties derive the whole of their moral force from the moral force of those general duties. It may not quite be the case that, existentially, they are wholly derivative from general duties. We cannot always deduce from considerations of general duties alone who in particular should take it on themselves to discharge them; where the general principle leaves that question open, some further (independent, often largely arbitrary) "responsibility principle" is required to specify it. Still, on this account, special duties are *largely* if not wholly derivative from considerations of general duty.

The practical consequences of this finding are substantial. If special duties can be shown to derive the whole of their moral force from their connections to general duties, then they are susceptible to being overridden (at least at the margins, or in exceptional circumstances) by those more general considerations. In this way, it turns out that "our fellow countrymen" are not so very special after all. The same thing that makes us worry mainly about them should also make us worry, at least a little, about the rest of the world, too.

These arguments draw upon larger themes developed elsewhere.³⁵ Here I shall concentrate narrowly upon their specific application to the problem of our special duties toward compatriots. The strategy I shall pursue here is to start from the presumption that there are, at root, no distinct special duties but only general ones. I then proceed to show how implementing those general duties gives rise to special duties much like those we observe in the practice of international relations. And finally I shall show how those special duties arising from general duties are much more tightly circumscribed in their extended implications than are the special duties deriving from any of the other models.

A

Let us start, then, from the assumption that we all have certain general duties, of both a positive and negative sort, toward one another. Those general injunctions get applied to specific people in a variety of ways. Some are quasi-naturalistic. Others are frankly social in character.

For an example of the former, suppose we operate under some general injunction to save someone who is drowning, if you and you alone can do so. Suppose, further, that you happen to find yourself in such a position one day. Then that general injunction becomes a compelling commandment addressed specifically to you.

The same example is easily adapted to provide an instance of the second mode as well. Suppose, now, that there are hundreds of people on the beach watching the drowning swimmer flounder. None is conspicuously closer or conspicuously the stronger swimmer; none is related to the swimmer. In short, none is in any way "naturalistically" picked out as the appropriate person to help. If all of them tried to help simultaneously, however they would merely get in each other's way; the probable result of such a melee would be multiple drownings rather than the single one now in prospect. Let us suppose, finally, that there is one person who is not naturalistically but, rather, "socially" picked out as the person who should effect the rescue: the duly-appointed lifeguard.³⁶ In such a case, it is clearly that person upon whom the general duty of rescue devolves as a special duty.

Notice that it is not a matter of indifference whom we choose to

³⁵ Goodin 1985c. Pettit and Goodin 1986. See similarly Shue 1988.

³⁶ This, incidentally, provides another explanation for why we should appoint lifeguards for crowded but not uncrowded beaches. Besides its being more efficient to put the lifeguards where there are more people potentially in need of being rescued, it is also more important where there are many potential rescuers that someone be picked out as the "obvious" one to render assistance in an emergency.

vest with special responsibility for discharging our general moral duties. Obviously, some people would, for purely naturalistic reasons, make better lifeguards than others. It is for these naturalistic reasons that we appoint them to the position rather than appointing someone else. But their special responsibility in the matter derives wholly from the fact that they *were* appointed, and not at all from any facts about why they were appointed.

Should the appointed individuals prove incompetent, then of course it is perfectly proper for us to retract their commissions and appoint others in their places. If responsibility is allocated merely upon the bases here suggested, then its reallocation is always a live issue. But it is an issue to be taken up at another level, and in another forum.¹⁷ Absent such a thoroughgoing reconsideration of the allocation of responsibilities, it will almost always be better to let those who have been assigned responsibility get on with the job. In all but the most exceptional cases of clear and gross incompetence on the part of the appointed individual, it will clearly be better to get out of the way and let the duly appointed lifeguard have an unimpeded chance at pulling the drowning swimmer out of the water.

That seems to provide a good model for many of our so-called special duties. A great many general duties point to tasks that, for one reason or another, are pursued more effectively if they are subdivided and particular people are assigned special responsibility for particular portions of the task. Sometimes the reason this is so has to do with the advantage of specialization and division of labor. Other times it has to do with lumpiness in the information required to do a good job, and the limits on people's capacity for processing requisite quantities of information about a great many cases at once. And still other times it is because there is some process at work (the adversarial system in law, or the psychological processes at work in child development, e.g.) that presupposes that each person will have some particular advocate and champion.¹⁸ Whatever the reason, however, it is simply the case that our general duties toward people are sometimes more effectively discharged by assigning special responsibility for that matter to some particular agents. When that is the case, then that clearly is what should be done.¹⁹

¹⁷ Thus the ascription of "role responsibilities" takes on the same two-tier structure familiar from discussions of "indirect consequentialism" (Hare 1981, pp. 135-40, 201-5; Williams 1983).

¹⁸ Nagel 1978, p. 81. Williams 1981, chap. 1.

¹⁹ Assigning responsibility to some might have the effect of letting others off the hook too easily, refusing to come to the aid of someone in distress because it is someone else's job. That just points to the importance of backup responsibilities, discussed farther on.

Thus, hospital patients are better cared for by being assigned to particular doctors rather than having all the hospital's doctors devote one n th of their time to each of the hospital's n patients. Someone accused of a crime is better served, legally, by being assigned some particular advocate, rather than having a different attorney appear from the common pool of attorneys to represent him at each different court date.⁴⁰ Of course, some doctors are better than others, and some lawyers are better than others; so it is not a matter of indifference which one is handling your case. But any one is better than all at once.

B

National boundaries, I suggest, perform much the same function. The duties that states (or, more precisely, their officials) have vis-à-vis their own citizens are not in any deep sense special. At root, they are merely the general duties that everyone has toward everyone else worldwide. National boundaries simply visit upon those particular state agents special responsibility for discharging those general obligations vis-à-vis those individuals who happen to be their own citizens.⁴¹

Nothing in this argument claims that one's nationality is a matter of indifference. There are all sorts of reasons for wishing national boundaries to be drawn in such a way that you are lumped together with others "of your own kind"; these range from mundane considerations of the ease and efficiency of administration to deep psychological attachments and a sense of self that may thereby be promoted.⁴² My only point is that those are all considerations that bear on the drawing and redrawing of boundaries; they are not, in and of themselves, the source of special responsibilities toward people with those shared characteristics.⁴³

The elementary facts about international responsibilities set out in

⁴⁰ This is the "division of labor model" of the adversary system discussed by Wasserstrom (1975, p. 9; 1983, p. 30).

⁴¹ This seems broadly in line with Christian Wolff's ([1749] 1934) early analysis. He certainly believes we have special duties toward our own nation (sec. 135); but it is clear from his preface (secs. 9-15) that those special rights and duties are set in the context of, and derived from, a scheme to promote the greater common good of all nations as a whole. Several modern writers canvass this argument, not always altogether approvingly (Shue 1980, pp. 139-44; Frankena 1977, p. 81; cf. Hare 1981, pp. 201-2).

⁴² Sidgwick 1891, chap. 14. Barry 1983. MacIntyre 1984. Cf. Cottingham 1986, pp. 370-4.

⁴³ That is to say, if general duties would be better discharged by assigning special responsibilities to a group of people who enjoy helping one another, then we should so assign responsibilities - not because there is anything intrinsically good about enjoying helping those particular others, but merely because that is the best means to the intrinsic good of discharging general duties.

Section II can all be regarded as fair "first approximations" to the implications of this assigned responsibility model. States are assigned special responsibility for protecting and promoting the interests of those who are their citizens. Other states do them a *prima facie* wrong when they inflict injuries on their citizens; it is the *prima facie* duty of a state, acting on behalf of injured citizens, to demand redress. But ordinarily no state has any claim against other states for positive assistance in promoting its own citizens' interests. That is its own responsibility. Among its own citizens, however, it is perfectly proper that in discharging that responsibility the state should compel its citizens to comply with various schemes that require occasional sacrifices so that all may prosper.⁴⁴

C

So far, the story is strictly analogous in its practical implications to that told about mutual-benefit societies in Section IV. Here, as there, we have special duties for promoting the well-being of compatriots. Here, as there, we are basically obliged to leave foreigners as we found them. The rationale is different. Here, it is that we have been assigned responsibility for compatriots in a way that we have not been assigned any responsibility for foreigners. But the end result is much the same – so far, at least.

There are, however, two important points of distinction between these stories. The first concerns the proper treatment of the useless and the helpless. So far as a mutual-benefit society is concerned, useless members would be superfluous members. Not only may they be cast out, they ought to be cast out. If the *raison d'être* of the society is mutual benefit, and those people are not benefiting anyone, then it is actually wrong, on mutual-benefit logic, for them to be included. (That is true, at least insofar as their inclusion is in any way costly to the rest of the society – ergo, it is clearly wrong, in those terms, for the severely handicapped to draw any benefits from a mutual-benefit society.) The same is true with the helpless, that is, refugees and stateless persons. If they are going to benefit society, then a mutual-benefit society ought to take them in. But if they are only going to be a net

⁴⁴ If example *g* in Section II is construed as a special positive duty toward aliens, then it poses something of a problem for all three other models of special responsibilities. All those would, for different reasons, expect positive duties to be stronger *vis-à-vis* compatriots, not toward aliens. The assigned responsibility model alone is capable of explaining the phenomenon, as a manifestation of our general duty toward everyone at large which persists even after special responsibilities have been allocated. More will be said of that residual general duty later.

drain on society (as most of the "boat people" presumably appeared to be, e.g.), then a mutual-benefit society not only may but *must*, on its own principles, deny them entry. The fact that they are without any other protector in the international system is, for mutual-benefit logic, neither here nor there.

My model, wherein states' special responsibilities are derived from general ones of everyone to everyone, cancels both those implications. States are stuck with the charges assigned to them, whether those people are a net benefit to the rest of society or not. Casting off useless members of society would simply amount to shirking their assigned responsibility.

The "helpless" constitute the converse case. They have been (or anyway they are now) assigned to no one particular state for protection. That does not mean that all states may therefore ignore or abuse them, however. Quite the contrary – what justifies states in pressing the particular claims of their own citizens is, on my account, the presumption that everyone has been assigned an advocate/protector.⁴⁵ Then, and only then, will a system of universal special pleading lead to maximal fulfillment of everyone's general duties toward everyone else worldwide.

Suppose, however, that someone has been left without a protector. Either he has never been assigned one, or else the one he was assigned has proved unwilling or unable to provide the sort of protection it was his job to provide. Then, far from being at the mercy of everyone, the person becomes the "residual responsibility" of all.⁴⁶ The situation here is akin to that of a hospital patient who, through some clerical error, was admitted with some acute illness without being assigned to any particular physician's list. He then becomes the residual responsibility of all staff physicians of that hospital.

To be sure, that responsibility is an "imperfect" one as against any particular state. It is the responsibility of the set of states, taken as a whole, to give the refugee a home; but it is not the duty of any one of them in particular.⁴⁷ At the very least, though, we can say this much. It would be wrong for any state to press the claims of its own citizens strongly, to the disadvantage of those who have no advocate

⁴⁵ Thus, international law gives aliens no right to protest directly to host states if they have been mistreated by it; instead they are expected to petition their home governments to make representations to the host state on their behalf (Oppenheim 1955, vol. 1, chap. 3).

⁴⁶ Goodin 1985c, chap. 5. Pettit and Goodin 1986, pp. 673–6.

⁴⁷ Shacknove 1985. Vattel [1758] 1863, bk. 1, chap. 19, sec. 230. See similarly Wolff [1749] 1934, secs. 147–9, and Grotius [1625] 1925, bk. 2, chap. 2, sec. 16. Vattel and Wolff specifically assert the right of the exile to dwell anywhere in the world that will have him, saying that hosts may properly refuse only for "good" and "special reasons."

in the system;⁴⁸ and it would not be wrong (as, perversely, it would be on the mutual-benefit–society model) for any state to agree to give refugees a home. Both these things follow from the fact that the state's special responsibility to its own citizens is, at root, derived from the same considerations that underlie its general duty to the refugee.

The second important difference between my model and mutual-benefit logic concerns the critique of international boundaries and the obligation to share resources between nations. On mutual-benefit logic, boundaries should circumscribe groups of people who produce benefits for one another. Expanding those boundaries is permissible only if by so doing we can incorporate yet more mutually beneficial collaborators into our society; contracting those boundaries is proper if by so doing we can expel some people who are nothing but liabilities to our cooperative unit. On mutual-benefit logic, furthermore, transfers across international boundaries are permissible only if they constitute mutually beneficial exchanges. The practical consequence of all this is, characteristically, that the rich get richer and the poor get poorer.

On the model I have proposed, none of this would follow. Special responsibilities are, on my account, assigned merely as an administrative device for discharging our general duties more efficiently. If that is the aim, then they should be assigned to agents capable of discharging them effectively; and that, in turn, means that sufficient resources ought to have been given to every such state agent to allow for the effective discharge of those responsibilities. If there has been a misallocation of some sort, so that some states have been assigned care of many more people than they have been assigned resources to care for them, then a reallocation is called for.⁴⁹ This follows not from any special theory of justice, but, rather, merely from the basis of special duties in general ones.⁵⁰

If some states prove incapable of discharging their responsibilities effectively, then they should either be reconstituted or assisted.⁵¹

⁴⁸ Similarly, in the "advocacy model" of the law, it is morally proper for attorneys to press their clients' cases as hard as they can if and only if everyone has legal representation. If institutions fail to guarantee that, it is wrong for attorneys to do so (Wasserstrom 1975, pp. 12–13; 1983).

⁴⁹ Miller (1988) rightly objects to putting the poorly-off in charge of the poorly-off, and the well-off in charge of the well-off. That is not a critique of my model, however, but rather a critique of existing international boundaries from within my model.

⁵⁰ Cf. Barry 1983.

⁵¹ Perhaps to the point of a poor state offering itself up as a colony of a richer one – but only insofar as the colonized have no strong interests in collective autonomy, and only insofar as the colonizers really would promote the interests of

Whereas on mutual-benefit logic it would actually be wrong for nations to take on burdens that would in no way benefit their citizens, on my model it would certainly not be wrong for them to do so; and it would in some diffuse way be right for them to do so, in discharge of the general duties that all of them share and that underwrite their own grant of special responsibility for their own citizens in the first place.⁵⁴

VI. CONCLUSION

Boundaries matter, I conclude. But it is the boundaries around people, not the boundaries around territories, that really matter morally. Territorial boundaries are merely useful devices for "matching" one person to one protector. Citizenship is merely a device for fixing special responsibility in some agent for discharging our general duties vis-à-vis each particular person. At root, however, it is the person and the general duty that we all have toward him that matters morally.

If all has gone well with the assignment of responsibilities, then respecting special responsibilities and the priority of compatriots to which they give rise would be the best way of discharging those general duties. But the assignment of responsibility will never work perfectly, and there is much to make us suppose that the assignment embodied in the present world system is very imperfect indeed. In such cases, the derivative special responsibilities cannot bar the way to our discharging the more general duties from which they are derived. In the present world system, it is often – perhaps ordinarily – wrong to give priority to the claims of our compatriots.⁵⁵

the colonized poor rather than exploit them. Neither condition, of course, is often likely to be satisfied.

⁵⁴ This duty to render assistance across poorly constituted boundaries might be seen as a "secondary, back-up responsibility" that comes into play when those assigned primary responsibility prove unwilling or unable to discharge it (Goodin 1985c, chap. 5).

⁵⁵ Earlier versions were presented to an ECPR Workshop in Amsterdam and to seminars at the Universities of Essex and Stockholm. I am grateful to all those audiences, and to Hillel Steiner, for comments.

Nuclear disarmament as a moral certainty

One of the most biting comments on postwar American history comes from the pen of Arthur Schlesinger, Jr. It goes like this: "As the traumatic experience of the Great Depression led to the resolution to make the economy depression-proof, so the traumatic experience of Pearl Harbor led to the resolution to make the nation war-proof." Laudable though those goals might be, Schlesinger's advice is simply, "Let's not get carried away" seeking certainty where none is available.¹

According to the now conventional wisdom, American foreign policy-makers in particular must "accept the fact of uncertainty and learn to live with it. Since no magic will provide certainty, our plans must work without it."² Modern deterrence theory is especially deeply imbued with this renunciation of the quest for certainties. There is no such thing as a perfect deterrent, we are firmly told. There is no absolute guarantee of perpetual peace. The Harvard Nuclear Study Group advises, "Any form of atomic escapism is a dead end. Living with nuclear weapons is our only hope. It requires that we persevere in reducing the likelihood of war even though we cannot remove the possibility altogether." Or in the immortal words of John Foster Dulles, "You have to take chances for peace, just as you must take chances in war."³

At this point, however, I can only echo Schlesinger's admonition. Let's not get carried away, renouncing the quest for certainties even where they can reasonably be sought. Even if we are not able to tie down everything, we can still hope to tie down some of the most important things. That, as I understand it, is what the movement for nuclear disarmament is all about. Its aim is to produce modal changes

¹ Schlesinger 1981.

² Wohlstetter 1962, p. 401.

³ Harvard Nuclear Study Group 1983, pp. 254-5. Dulles is quoted in Ellsberg 1968, p. 38.

in the possibility of a large-scale nuclear war rather than just marginal changes in its probability. That limited quest for certainties is, I shall argue, eminently defensible.

I

To understand what I am arguing for, it might help to know what I am arguing against. That is nothing less than nuclear deterrence itself. Its basic idea is that each side can, by emphasizing the risks of nuclear war, deter the other from steps that might turn that risk into a reality. But this is very much a gamble since for the deterrent to be credible the risk must be real. So in essence deterrence is a scheme for making nuclear war less probable by making it more probable.⁴

That the logic of nuclear deterrence is incorrigibly probabilistic is widely acknowledged. Strategists and social scientists describe as its "fundamental premise" the proposition that "nuclear weapons make *probable* the rejection [by one's opponents] of armed aggression as a potential policy alternative" (emphasis added). Moral philosophers also fully acknowledge the nature of the gamble. In the words of Bernard Williams, "The morality of deterrence is, I think, legitimately one in which you think principally about those steps which make it *less likely* that the weapons get used" (emphasis added).⁵

My argument is that all such notions of probability and likelihood are simply inappropriate in these circumstances. Maybe such concepts are not even meaningful at all when applied to situations involving reflective human agents rather than mere random processes.⁶ But in any case it is clear that, where probabilities of nuclear war are concerned, we just do not know enough about the shape of the underlying distribution to justify employing any of the standard techniques for estimating probabilities. That we can have no reliable probability estimates is in itself quite enough to render probabilistic reasoning about such affairs wildly inappropriate.

Certainly we have no solid objective statistics, based on frequency counts or such like. The balance of terror has kept the peace for the past thirty-five years, to be sure. But thirty-five years is just too short

⁴ Literally, of course, this is a non sequitur, but the phrase captures something of the spirit of deterrence. Note, for example, the Harvard Nuclear Study Group's (1983, p. 15, emphasis added) talk of "taking a *small risk* of nuclear war" in hopes of "preventing an even *larger risk* of nuclear war," or Blechman and Hart's (1982) discussion of how the "threat to lose control" can serve as a deterrent threat. See similarly Paskins and Doctrill 1979, pp. 64-77, 239, and Trachtenberg 1985, pp. 359-61.

⁵ Brody 1968, p. 130. B. Williams 1982, pp. 289-90. See similarly Gauthier 1984.

⁶ Burns 1970.

a run on which to base our probability judgments, given the unacceptability of even very small probabilities of such a very great horror. Besides, nuclear war is just not the sort of thing whose probabilities we dare to estimate by trial-and-error procedures – the first error may well mark the learner's own end.⁷

Nor do we have any well-validated scientific theories (about, e.g., the genesis and escalation of international conflicts) from which we might hope to derive reliable estimates of the probability of a breakdown in deterrence which would lead to a large-scale nuclear war. We suffer not from a lack of such theories but rather from a surfeit of them; and none can prove itself decisively superior to all the others.⁸

Nor, finally, do we have any particularly good reason to place any great faith in subjective probability estimates. Of course, we can always bully people into stating their "best guess" as to the chances of anything occurring; we can even bully them into rendering those probability estimates consistent. But when such estimates are as groundless as those concerning the chances of nuclear deterrence collapsing into nuclear war are, we should not set any great store by them. Ellsberg says, "It's no use bullying me into taking action . . . by flattering my 'best judgment.' I know how little that's based on."⁹ Alas, most people do not. Psychological evidence suggests not only that "individuals are poor probability assessors" but also, "and perhaps more important, that they underestimate their poorness by assessing probabilities too tightly."¹⁰ Knowing this – and knowing all the severe distortions to which judgments under uncertainty are prone¹¹ – it would be sheer folly for us to predicate any profoundly important policy choices on such fallible subjective probability estimates.

The upshot is that it is altogether inappropriate to engage in probabilistic reasoning about the chances of a breakdown in the balance of terror that leads to a large-scale nuclear war. Objective statistics are unavailable; theories are too numerous and too divergent; subjective estimates are known to be too unreliable.¹² The problem is not just

⁷ All the standard arguments about nuclear reactor safety apply here. See Goodin 1982b, chap. 10.

⁸ "Where we have several competing theories which give different predictions, all these theories should be regarded with suspicion, and we should be prepared for a risk that is much higher than what is predicted by any of the theories" (Føllesdal 1979, p. 405). For a critical survey of competing theories, see Hoole and Zinnes 1976.

⁹ Ellsberg 1961, p. 663. See similarly Luce and Raiffa 1957, p. 304.

¹⁰ Zeckhauser 1975, p. 445. For evidence, see Kahneman et al. 1982, pp. 287–354.

¹¹ See Kahneman et al. 1982 in general and, more particularly, Fischhoff 1983.

¹² See similarly Elster 1979a, Page 1978, pp. 225–9, Kavka 1980 and Benn 1983.

that we cannot estimate point probabilities with any great precision – that we cannot say whether the probability of nuclear war this century is 10 percent or 15 percent. Nor is it even just that we cannot make the sorts of order-of-magnitude judgments that would allow us to make ordinal judgments about relative probabilities. We are in a worse situation still. We cannot even say with confidence in what direction any particular strategic innovation pushes the probability of all-out nuclear war. Some theories maintain that that risk is increased by cruise missiles or space-based defenses or nuclear proliferation. Others hold the opposite.¹¹ Neither logic nor experience enables us to choose confidently between these theories, and only a fool would trust unaided hunches with so much at stake.

The most that can be claimed for deterrence is that it will probably work to prevent war. So if probabilistic reasoning is inappropriate in these circumstances, deterrence is too. In short, my complaint against nuclear deterrence is that it amounts to playing the odds without knowing the odds. That constitutes recklessness par excellence. It would be the height of irresponsibility for anyone to wager the family home on rolls of such radically unpredictable dice. Where millions of lives are at stake, that judgment must surely apply even more harshly.

The conventional wisdom holds that the only responsible response to such radical uncertainty would be to adopt a diversified defense portfolio, hedging against all the conceivable risks. But of course it is in the nature of such things that, in the course of spreading ourselves to secure some protection against all possibilities, we leave ourselves less than fully protected against any of them. And, furthermore, some of the things that diversifying our defense portfolio might dictate (e.g., damage-limitation or postwar-reconstruction planning) could be seen as provocative by our opponents, thereby increasing the risk of the very worst eventualities we are hoping to protect against. Risk hedging is not a bad strategy if it is the best we can do. Often it is. Sometimes, however, another strategy is available, and when it is, it proves decidedly superior. It is to this new strategy that I now turn.

II

Here I shall offer a method of approaching such problems that deals in possibilistic rather than probabilistic terms. In possibilistic reasoning, there are only three categories to concern us: (1) the impossible,

¹¹ See, e.g., the debate between Payne and Gray (1984). For a novel argument that nuclear proliferation might reduce the chances of nuclear war by making all of us more careful, see Wildavsky 1962.

(2) the possible, and (3) the certain.¹⁴ And there are only four logically distinct ways of changing possibilities:

1. something previously impossible can be made possible;
2. something previously merely possible can be made certain;

or moving in the opposite direction:

3. something previously certain can be made merely possible; or
4. something previously possible can be made impossible.¹⁵

Some special significance of a logical sort clearly attaches to movements across the boundaries that define these modal categories. Where the outcomes in question are themselves matters of moral concern, that special significance is moral as well. If an outcome would be morally desirable, then (other things being equal) it is morally desirable for that outcome to be made not just more probable but certain or for that outcome to be rendered possible if before it was not. Conversely, if an outcome would be morally undesirable, then (other things being equal) it is morally desirable for that outcome to be made not just less probable but impossible or for it to be reduced to a mere possibility if before it was a certainty.

For those steeped in modern Bayesian decision theory, replacing probabilistic with possibilistic reasoning might seem queer indeed. For a wide variety of people in a wide variety of practical circumstances, however, it seems to come very naturally. Of course, that does not prove that there are good reasons for their so acting – arguments to that effect will come later. But as an important preliminary to those arguments, let me first respond to the “argument from queerness” by illustrating the role of possibilistic reasoning in everyday affairs. Consider first the evidence of a wide variety of psychological experiments showing that there is something like a “certainty effect” governing people’s responses to choice situations. However they choose between ordinary gambles, people seem disproportionately sensitive to outcomes that are certain. Even by the standards set by their own previous choices among gambles, people will pay considerably more than they should (in terms of probabilistically expected costs) to avoid a loss that is certain or (in terms of probabilistically expected gains foregone) to secure a gain that is certain.¹⁶

The same sort of logic seems to pervade the White House in crisis

¹⁴ Elster 1978. Hamblin 1959.

¹⁵ For completeness, we might add two further composite categories: (5) something previously impossible can be made certain (combining 1 and 2); or (6) something previously certain can be made impossible (combining 3 and 4).

¹⁶ Tversky 1975. Kahneman and Tversky 1979. This grows out of work on the Allais Paradox (Luce and Raiffa 1957, sec. 13.5; Ellsberg 1961; Allais and Hagen 1979).

decision making. There, too, the emphasis seems to fall heavily on certainties and impossibilities rather than on fine-grained probability estimates. Notice that, during the Cuban missile crisis, President Kennedy justified taking a risk of nuclear war on the grounds that it was necessary "to avoid *certain* war later." Robert Kennedy added that American actions did not cross any modal boundaries on the other side either, saying that "we all agreed in the end that if the Russians were ready to go to nuclear war over Cuba, they were ready to go to nuclear war, and that was that. So we might as well have the show-down then as six months later."¹⁷ Whether or not any of this was true – or whether or not it was even genuinely thought to be true, at the time – it is nonetheless revealing that the rationalizations (if that is what they are) are phrased in terms of modalities, the certainty of war, and such like.

The same emphasis on possibilistic reasoning is also reflected in the special importance traditionally attached, both legally and morally, to the "first intervening agent" and the "last intervening agent" in any causal chain. Both make essentially possibilistic contributions to the outcome. The intervention of the first makes it possible, that of the last inevitable. It is hard to say precisely how to apportion responsibility between these two agents.¹⁸ Between them, however, the first and last intervening agents were traditionally thought to bear the bulk of the responsibility for the outcome. Recently we have come to appreciate the importance of other agents along the way – those who could have put an end to the causal chain but who failed to do so. In law, the increasingly popular doctrine of "comparative negligence" shares out some of the blame among them, also.¹⁹ What is significant in the present context, however, is that first, last, and intermediate agents are all blamed for their possibilistic contributions to the causal chain. Just as the first is blamed for making the harmful outcome possible and the last for allowing it to become inevitable, intermediate agents are blamed for failing to take opportunities when they had them to make the harmful outcome impossible.

All this is merely by way of refuting the "queerness" objection.

¹⁷ Snyder 1978, pp. 357, 345. See also Jervis 1979, pp. 310–11.

¹⁸ Prior 1956. Raphael 1956. Hart and Honoré [1959] 1985. Prosser 1971, chap. 12.

¹⁹ Prosser (1971, chap. 12) enters a forceful plea for "comparative negligence." Notice that the older, harsher doctrine of "contributory negligence" only makes sense if the plaintiff could have done something to prevent the harm from befalling himself; likewise, "vicarious liability" only makes sense on the assumption that masters/parents/owners could have done something to prevent their servants/children/dogs from inflicting the harm. Unless they could have done something to make the harmful outcome impossible – i.e., unless their negligence was a necessary condition of that outcome – it makes no sense to assign them liability for it.

Be the intuitions of Bayesians as they may, possibilistic considerations certainly do seem to loom large in many areas of ordinary decision making. Now I want to turn to argue that there is a class of extraordinary decision situations in which possibilistic considerations deserve special and, indeed, exclusive emphasis. In standard applications, possibilistic notions rightly operate in conjunction with probabilistic ones. Whether and how much we praise or blame people for their actions ordinarily depends at least in part on how probable it was (or is) that the possibilities they opened up would be actualized.²⁰

Under certain circumstances, however, those probabilistic elements drop out of the calculations, leaving us with a purely possibilistic analysis of moral responsibility. Here I shall concentrate on two such conditions. These are not necessarily mutually exclusive; indeed, both arguably obtain in the case of nuclear weapons decision-making. Nor are they necessarily exhaustive of all the circumstances that force us to rely on purely possibilistic reasoning. Still, they seem to be the most important ones for the present application at least.

One relatively familiar factor that forces us to shy away from probabilistic reasoning is the magnitude of the possible payoffs. Where there is any risk of something infinitely awful happening, then probabilities simply do not matter. Just so long as that outcome is possible – so long as there is some nonzero probability of its happening – we must do whatever we can to avoid it.²¹ Infinite costs, discounted by any probability larger than zero, are still infinite. And while few payoffs are literally infinite, the same sort of argument might still work where possible payoffs are virtually infinite. For all intents and purposes, probabilities might still drop out of our reckonings for much the same reason as before.²²

All that is, I think, both true and important. But it is also pretty standard. My more novel suggestion is that there is a second condition, completely independent of the first, which might also force us to

²⁰ In tort law, e.g., people are not held responsible for unforeseeable “freak” accidents. Some might analyze that as “anything . . . which [the actor] could not fairly have been expected to contemplate as possible, and therefore to avoid” (Holmes 1881, p. 94). The more standard analysis analyzes that notion in terms of what could be foreseen as both “possible and probable” (Hart and Honore [1959] 1985, chap. 9).

²¹ The most plausible way anything could have infinite disutility would be by foreclosing an otherwise potentially infinite stream of future payoffs (Dryzek 1983). In the limiting case there is a “zero-infinity dilemma,” i.e., “a virtually zero probability of a virtually infinite catastrophe” (Page 1978, pp. 208–12).

²² Similarly, tort law assigns anyone engaged in “ultrahazardous activities” (or “abnormally dangerous activities,” as they are now called) strict liability for any harms that might result, blatantly ignoring the probabilistic calculus that characterizes ordinary standards of negligence (Prosser 1965, chap. 21).

fall back on possibilistic rather than probabilistic standards of moral responsibility. That condition is as follows:

1. the agent's choice among alternative courses of action might make a morally significant difference in the outcome;
2. the agent's judgment about the probability that any particular outcome will result from any particular action is (and is or should be known by the agent to be) highly unreliable; and
3. the agent's judgment about what outcomes his various actions would make possible or impossible is (and is or should be known by the agent to be) highly reliable.⁴³

The emphasis in this second argument is on the differential reliability of possibilistic judgments over probabilistic ones. Often there will be no such differential. Whatever makes us uncertain of probabilities also makes us uncertain about possibilities. But at least occasionally we will have good grounds – which almost inevitably have to be good theoretical grounds rather than experiential ones – for supposing that some courses of action will make an outcome impossible, whereas other courses of action merely alter its probability in ways we cannot reliably predict.

Under this perhaps peculiar constellation of circumstances identified by my second argument for possibilistic reasoning, the only way an agent can confidently exert effective control over morally important consequences is by manipulating possibilities. And if the consequences really do matter morally, then clearly he should do so. This means that, *ceteris paribus*, a moral agent should make morally desirable outcomes possible or certain (and morally undesirable ones impossible or uncertain) wherever he can; that a moral agent should open up as many possible paths to good outcomes (and close off as many possible paths to bad ones) as he can; *et cetera*.

Under either of these two sets of circumstances, the morally responsible course of behavior is to pursue policies producing modal changes in the desired direction. The first argument tells us that we should aim, above all else, to remove the certainty or to guarantee the impossibility of outcomes that are really heinous. The second tells us that, even where the payoffs are less dramatic, we should nonetheless concentrate on possibilities if probability judgments are deemed grossly unreliable. Then all that is left is for us to choose between alternative courses of action on the basis of the possibilities they open up or close off for good or bad outcomes. Either of these conditions

⁴³ How reliable "reliable enough" is (and for that matter, how certain "certain enough" is) for the purposes at hand is a function of potential costs and benefits. It is permissible to play unreliable odds for small stakes, in a way it is not for large ones (Ellsberg 1961, p. 663; Luce and Raiffa 1957, p. 27).

alone would force us to shift from probabilistic to possibilistic reasoning. Where both conditions are operative, as I shall shortly argue they are in the case of nuclear armaments, they can be powerfully reinforcing.

This approach contrasts sharply both with the standard ethical practice of assessing policies in terms of their probable consequences and with the standard deontological alternative to that approach. Unlike the former, I shun probabilities; unlike the latter, I still want to embrace consequences.⁴⁴ My approach also contrasts sharply with all the standard decision-theoretic rules that would allow us to dispense with reckonings of probabilities.⁴⁵ Maximin, maximax, minimax regret, and Arrow-Hurwicz rules all choose between options on the basis of possible payoffs, without asking how probable they might be. But all these rules are preoccupied with limiting cases – with the best or worst or most regrettable possible outcome. Thus they ignore the existence of a great many other possibilities that might be almost as bad or as good; they likewise ignore how many or how few possible paths might be open to each of these outcomes. In short, decision theorists, in their zeal to break loose of the influence of probabilistic reasoning, have come to neglect important changes even in the set of possibilities facing us. Remedying that neglect is one of the major tasks of this chapter.

III

This focus on the moral importance of modal shifts can be shown to have important implications for nuclear weapons policy. The preconditions for applying my argument surely all exist. Little need be said to justify the claim that the consequences in view matter morally. Maybe consequentialistic considerations are not the only ones that should guide our choices, of military policies or any others; but where the consequences in view are so momentous as those involved in an all-out nuclear war, it would be sheer lunacy to deny such considerations any role at all.⁴⁶

For present purposes, there is no need to specify what makes the

⁴⁴ Benn (1983) leaps from the observation that probabilities are unreliable where nuclear strategies are concerned to the conclusion that consequentialistic reasoning is therefore precluded, thereby overlooking the option (possibilistic consequentialism) for which I am here arguing.

⁴⁵ Luce and Raiffa 1957, chap. 13. Lackey 1982; 1984.

⁴⁶ "It is thoroughly unreasonable to suppose that the goodness or badness of an action is entirely independent of its probable consequences, and no one but a moralist riding a theory to death would maintain this view for an instant" (Broad 1916, p. 278; see similarly B. Williams 1973a, p. 90).

consequences of a large-scale detonation of nuclear weapons so morally heinous. We can leave open the question of whether it is the dead bodies (or cities or civilizations) that should offend us or whether it is the violations of moral rights and duties that would inevitably be entailed by a full-blown nuclear war. The question here is merely how to allocate responsibility for avoiding outcomes that can be seen to be morally heinous from any number of perspectives.

Nor is there much doubt about the existence of both conditions independently driving us from probabilistic principles to possibilistic ones. As for the first, the potential costs of an all-out nuclear war surely are enormous and surely must count as "virtually infinite" if any do. As for the second, my introductory argument has already shown that, where risks of nuclear war are concerned, probabilistic reasoning is inappropriate.

What crucially remains to be shown is that there is any action which we can confidently predict *will* make a modal change in the desired direction. Often it seems we know no more about possibilities than about probabilities, or that opening up some bad possibilities is the necessary price of closing off others equally bad, or that opening up some bad possibilities is the necessary price of opening up some others that are equally good. This sort of indeterminacy, at least, seems absent in the case of nuclear weapons policy.

Consider first the responsibilities involved in moving up the ladder of possibility, from impossibility to possibility to certainty. According to the above analysis, there would be two classes of prime candidates. One would be those who were responsible for creating (inventing, funding and building) the nuclear weapons in the first place – for making nuclear war a possibility. The other would be those who would be responsible for taking the last step that makes all-out nuclear war inevitable. One of the great troubles with this as with all great "accidents" is that it is often not clear, either in retrospect or, much less, in prospect, what the "last step" is before some inexorable chain of events is set in motion. But in the case of nuclear war, this much at least is certain. The first superpower to launch nuclear weapons directed at the other's homeland has thereby made all-out nuclear war inevitable.

Some might say that nothing is "inevitable," in any morally important sense, so long as other people retain some subsequent choice in the matter. Thus when you refuse a blackmailer's demands, knowing full well that he *will* kill the child, that does not burden you with the responsibility for the child's death.²⁷ Nor on this account would the agent launching the first nuclear strike be responsible for the en-

²⁷ Davis 1980a, pp. 201–10; 1980b, pp. 31–4.

suing exchange of weapons – if the other side could, in some meaningful sense, choose whether to launch a second strike. But there is an equally persuasive argument in the opposite direction. If you could have foreseen the other's reaction, then you should have taken account of it in making your own decision; if you have failed to do so and the injury occurs, then both of you are liable for it.²⁸

Besides, in the case of nuclear war there are powerful reasons for doubting that the other side really could "choose" in any meaningful sense whether or not to launch a retaliatory strike. One reason is purely psychological – turning the other cheek after an unprovoked nuclear attack may well be beyond the psychological capacities of most people. Another even more important reason is technological. Most strategists now seem to agree that a first strike would try (probably successfully) to knock out some important components of the other side's command-control-communications systems.²⁹ That means that the counterstrike would be fired by a headless automaton. The official line, and the fond hope, is that no one cut off from the commander in chief will have authority to launch a nuclear counterstrike. The great danger is that no one, in those circumstances, will have the effective authority to stop it.

At this point, nuclear weapons have been discovered, built and deployed. Those who did so may have a lot to answer for. But from the perspective of present policy-making, all that is history. The question before us is what, given that history, we are now to do. To present policy-makers therefore falls a peculiarly strong responsibility to make sure that, now that nuclear war is possible, they do nothing to make it inevitable. That, as I have just argued, means that each side should abstain from any first nuclear strike on the other superpower's homeland.

Next let us consider the responsibilities involved in moving down the ladder of possibility, from certainty to possibility to impossibility. Again, there are two classes of prime candidates for responsibility. One would be those who were responsible for averting an otherwise certain nuclear exchange – for transforming the certainty of nuclear war into the mere possibility of one. The other would be those who were responsible for making an all-out nuclear war impossible.

It is difficult to imagine what desperate circumstances might be described by the first category (an all-out nuclear war being inevitable in the absence of the agent's intervention). But suppose, for example, that the troops in the silos (theirs or ours) had standing orders to launch a nuclear attack under certain narrowly specified conditions,

²⁸ Beale 1895, p. 87.

²⁹ Ball 1981. Keeny and Panofsky 1981–2. Bracken 1983.

unless instructed otherwise by their commander in chief.³⁰ Then an all-out nuclear war would be inevitable unless the commander in chief gives the order for them to hold their fire. He would, therefore, have a peculiarly strong responsibility to do so. Much the same can be said for everyone along the chain of command that would be responsible for transmitting the commander in chief's order to launch a nuclear attack or counterattack. All-out nuclear war would be inevitable unless one or some group of them acts so as to prevent that message from getting through. Consequently each of them would have a peculiarly strong responsibility so to act.

The second category is, however, of more immediate relevance. Whoever has an opportunity to make all-out nuclear war impossible would, on this argument, have a peculiarly strong responsibility to seize that opportunity. Herein lies the case for nuclear disarmament. Of course, if all sides were to undertake such a policy, nuclear war between them would be quite literally impossible. Universal nuclear disarmament is most strongly to be commended on these grounds, therefore. But under certain circumstances – circumstances which arguably obtain in the present world – unilateral nuclear disarmament by one of the two superpowers, combined with a reorientation of its conventional force posture, would produce the same effect. There is simply no credible scenario by which a nuclear-armed superpower – provided it is at once minimally rational and governed by the standard goals guiding world politics – would, either by accident or by design, be led to launch a full-scale nuclear assault on an opponent armed only with conventional weapons of a merely defensive sort.³¹

³⁰ These circumstances are, alas, not so fanciful as they might seem. It is widely reported that North American Air Defense (NORAD) commanders used to have authority, in certain circumstances, to launch a nuclear attack on their own authority. Certainly there were frequent demands for "preplanning" of nuclear strikes by NATO field commanders, incorporating the description of circumstances in which they are authorized to fire in the Single Integrated Operational Plan (Frei 1983; Pringle and Arkin 1983).

³¹ Marc Trachtenberg, in comments at the Aspen conference, argued that history belies this point: "Consider the example of the United States in its period of atomic monopoly. If war had broken out around 1950, this country, which thought of itself as liberal, civilized and humane, would almost certainly have dropped every bomb in its arsenal on the Soviet Union, and especially Soviet 'urban industrial' targets. If America would have behaved in this way, are we certain that the Soviets, given similar circumstances, would not?" (see similarly Freedman 1981, chap. 4). But the circumstances would not be similar. The American nuclear attack was contemplated only in the event of an overwhelming conventional assault by Soviet forces on America's European allies. Were a superpower to undertake both unilateral nuclear disarmament and the reorientation of conventional forces into a purely defensive posture, then neither history

In a war of conquest, no aggressor strives to destroy its spoils.¹²

This case for unilateral nuclear disarmament is hedged in various respects. Let me elaborate a little. Notice, first, that I am assuming a world in which there are only two nuclear superpowers. Only in that case would unilateral nuclear disarmament by one party preclude an all-out nuclear war. Mine is an argument for reducing the number of nuclear-armed superpowers to one, not just *by* one. Were there more than two vast nuclear arsenals in the world, the same sort of argument would lead to a call for unilateral nuclear disarmament, ideally by all nuclear superpowers but at the very least by all but one.¹³

Notice, second, that I merely claim that unilateral nuclear disarmament would avert "full-scale" or "all-out" nuclear war. Nuclear weapons might still be used in anger. There are still the lesser nuclear powers who might make some little mischief; and if one superpower retains nuclear weapons, it might occasionally detonate some of them, either to back up its blackmail or to prevent the other from rearming. Naturally, no nuclear blast is to be welcomed. Nonetheless, I assume that there are scale effects and that, for a variety of reasons, an all-out

nor strategic doctrine more generally yields any credible scenario by which it would be subject to an all-out nuclear assault.

"No credible scenario" is indeed the proper standard of impossibility. Possible worlds, recall, must be constructed out of possible steps from the actual world; and what is "incredible" is inconceivable and hence impossible, at least from our present perspective. There are of course "degrees of credibility," resolvable into an ordinary probability calculus. Here I focus instead on the limiting case of impossibility, i.e., the situation in which there is no scenario with any credibility at all leading to the outcome in question.

¹² Doug Lackey, in comments at the Aspen conference, offered as a *reductio* the argument that following this logic we would be obliged to remove all potentially lethal objects from our homes and thereby render ourselves incapable of committing murder. Although I certainly think it would be a good idea for Americans to remove some of the lethal weapons from their homes (e.g., handguns), I am happy enough for them to hang onto others (e.g., kitchen knives and frozen legs of lamb). The reason, as Lackey intimated, has to do with the relative probabilities that they will in fact be used to commit murders. Long experience with guns, carving knives and legs of lamb lying around the house grounds reasonable estimates of the probabilities that they will be used as murder weapons. Our all-too-brief experience with nuclear weapons, alas, offers no basis for reliable probability estimates.

¹³ Strategically, each nation will hold out, each wanting to be that one nuclear superpower left in an otherwise nuclear-disarmed world. The dangers attending such an unequal coordination game constitute the best argument for unilateral disarmament not only in the two-nation but also in the *n*-nation case (Hardin 1983, pp. 249-54). As one superpower divests itself of nuclear weapons, lesser powers that used to live under its nuclear umbrella may feel compelled to acquire or increase nuclear arsenals of their own, of course; but so long as those arsenals remain below some threshold (discussed in the next paragraph) "all-out nuclear war" would still be impossible.

nuclear war would be far, far worse than the detonation of just a few nuclear devices. Where this threshold lies depends, of course, on what precisely it is that makes an all-out nuclear war so much worse than a limited one. If it is the prospect of a "nuclear winter," and if (as some suggest) detonating even as few nuclear weapons as France will soon possess might trigger it, then once again we would need concerted action aimed at more widespread nuclear disarmament in order to guarantee that the threshold would never be crossed.¹⁴ I ignore that possibility here not because I think it unimportant but merely because if true it constitutes an even more powerful argument for the suicidal superfluousness of nuclear weapons than my own.

Third, notice that unilateral nuclear disarmament renders all-out nuclear war impossible only if the remaining nuclear superpower is at once "minimally rational" and in pursuit of "the standard goals guiding world politics." These caveats are necessary because a grossly irrational power might launch an all-out nuclear assault on a nuclear-disarmed opponent, however counterproductive that might be.¹⁵ And a nation bent on the pursuit not of wealth, status and power but, rather, of genocide would not regard it as counterproductive at all – the complete destruction of the loathed group, its artifacts and culture, is precisely the end in view. The only way of rendering all-out nuclear war impossible, when a mad/genocidal power is involved, is for that power to be deprived of a nuclear arsenal of any size itself. Certainly there can be no guarantee that we (assuming we are the ones who are neither mad nor genocidal) will not be subjected to a full-scale nuclear attack by such a power, even if we undertake unilateral nuclear disarmament ourselves. In the terms of my argument, that means that we would be morally at liberty to retain a nuclear arsenal ourselves. That, however, is not to say that so doing is morally obligatory or even necessarily advisable. (There is even less reason to believe that mad/genocidal opponents will be deterred by our nuclear arsenal or be induced by it to bargain with us for mutual arms reductions than there is to believe that ordinary opponents will respond in these ways; and, as I argued at the outset, that is not enough to bank on even in

¹⁴ Sagan 1983-4, p. 286.

¹⁵ Jan Narveson, at the Apsen conference, teased me with the paradox that, if I am presupposing really rational actors, then the best way to guarantee nuclear peace is through deterrence, since really rational actors would never fall into an all-out nuclear war. But that presupposes more rationality ("superrationality") than my argument requires. Much evidence suggests that superpowers, facing the perceived threat of imminent nuclear destruction, behave quite irrationally (Snyder 1978; Jervis 1979). All my condition of "minimal rationality" requires, however, is that a nuclear-armed superpower, facing no real threat to its strategic interests (specifically, no nuclear threat), should refrain from gratuitously destroying some other nation. Presumably that sort of rationality is common enough.

the ordinary case.) It is merely to say that the moral case for us engaging unilaterally in nuclear disarmament loses much of its force under such circumstances.

Thus, my case for unilateral nuclear disarmament depends on various empirical assumptions. I assume that there is only one other superpower besides the United States with a nuclear arsenal sufficiently large to produce whatever effects we fear from an all-out nuclear war and that that would remain true even after an American nuclear disarmament. Furthermore, I assume that, for now and the foreseeable future, none of the world's nuclear superpowers are either mad or genocidal. With these assumptions, unilateral nuclear disarmament combined with a strictly defensive conventional force posture on our part would render all-out nuclear war impossible; and such a policy is, therefore, morally mandatory on the argument I have here been discussing. Anyone wishing to reject this conclusion must find good grounds for rejecting either that moral theory or one of those empirical assumptions – and these must be grounds that do not lead to an even stronger indictment of nuclear weapons than my own.

Some might seize on my last pair of caveats for an excuse for maintaining a nuclear arsenal as a kind of hedge against the eventuality that a mad or genocidal opponent might someday emerge. That is, after all, a possibility – just the sort of possibility that has been plaguing me elsewhere in this essay. The difference, however, is this. We have long and sad experience of mad and genocidal rulers; and based on this experience, we have a pretty good basis for predicting them before they emerge or at the very least for picking them out long before they have a chance to make any real trouble for us. In short, I am inclined to assume that we can see mad/genocidal leaders coming and rearm in time to deal with them, if nuclear arms or offensive conventional forces are indeed the right way to deal with them at all.¹⁶

IV

So far I have been discussing the problem of how best to pursue a single goal. Nuclear deterrence, remember, was defended by the Harvard Nuclear Study Group as a matter of taking a smaller risk of nuclear war to avoid a larger risk of nuclear war.¹⁷ My conclusion has

¹⁶ Colin Gray, in his comments at the Aspen conference, replied that (1) we are not very good at predicting crazy leaders far in advance and (2) it may take a decade or more to reorient our defense posture completely. But no leader, however mad or however well armed, is in a position to initiate the holocaust on the first day in office. And while some defense plans would take a decade to implement, those are not the only (or even most likely) ones to deter a mad or genocidal opponent.

¹⁷ Harvard Nuclear Study Group 1983, p. 15.

been that, owing to certain peculiar features of the problem (which may in fact be common to a wide range of other problems), the only justifiable strategy for pursuing that goal is to seek changes in the possibility of nuclear war occurring rather than merely changes in its probability.

However, avoiding nuclear war is not the only goal that nuclear deterrence is meant to serve. Among other things, it is also meant to prevent a nation from being conquered by tyrants; and that goal is not extensionally equivalent to the first. (The rule of tyrants might be avoided by going to nuclear war, after all.) It is typical for a policy to impinge on a multiplicity of independent goals in this way. Therefore, the prescriptions for possibilistic evaluation of policy options, couched here in *ceteris paribus* terms for the analysis of impacts on goals one by one, must be broadened so as to address situations in which one goal must be traded off for another.

I have little novel to say on the broader questions of how, whether and when we should make trade-offs between moral goals. Here I shall confine myself merely to emphasizing that, even in a world of moral trade-offs, shifting from probabilistic to possibilistic reasoning still does make some practical difference.

Talk of trade-offs conjures up various images that fit comfortably with the probabilistic model. We ordinarily – and ordinarily rightly – think in terms of swapping a little more of this for a little less of that or a little greater risk of one evil for a little less of another. But if probabilistic reasoning has been deemed inappropriate where one of the goals (e.g., avoiding all-out nuclear war) is concerned, then there can be no question of swapping a little more risk of that for a little less of any other evil. In the context of that goal, “more” and “less” risk are meaningless terms. We are obliged to think instead in possibilistic terms and trade off that goal (if we do so at all) only in very large modal lumps. We can only ask how much it is worth (in terms of some other goal, such as avoiding tyranny) to have nuclear war made possible where it might have been made impossible or how much it is worth to have it made certain where it might have been merely possible. We cannot talk meaningfully in terms of incremental trades within those large modal lumps.

Of course, none of that is to deny that we should trade large lumps of one goal for some quantity (large or small) of some other goal. It all depends on the relative importance of each goal. We rightly trade away large lumps of relatively unimportant goals all the time. But the more important the goal, the more inclined we are to opt for the “all” branch of the “all-or-nothing” choice that such lumpiness forces on us.

This has some new and important implications for nuclear policy

debates. Traditionally, ban-the-bombers have felt obliged to argue the immorality of making any trade-off at all between the goal of avoiding all-out nuclear war and any other goals, such as (merely) avoiding tyranny. Or, at the very least, they felt obliged to argue that the former was morally so terribly much more important that, even if trade-offs were allowed in principle, the latter goals would never have any real chance of trumping the former.

Both those propositions seem to me to be defensible, and from a surprising range of moral perspectives.¹⁶ If the preceding arguments are correct, however, both claims are far stronger than strictly necessary in order for the goal of avoiding all-out nuclear war to prevail in this trade-off. All that is strictly necessary is that that goal be at least as important, morally, as all the countervailing moral considerations taken together. Then, given the peculiarities of the possibilities facing us, the goal of avoiding all-out nuclear war will automatically prevail.

The argument goes like this. The modal lumpiness inherent in the goal of preventing all-out nuclear war means that we must pursue that goal in an all-or-nothing fashion. If we opt for an all-out pursuit of that goal, then we can have that and more since there would still be some things we could do to promote – to some extent or another – those other goals. (Conventional deterrents or passive defenses may be more-or-less effective ways of avoiding rule by tyrants, even if we dismantle unilaterally our nuclear arsenal.)¹⁷ If we opt instead not to pursue the goal of avoiding all-out nuclear war at all, then the most we could possibly hope to accomplish would be complete satisfaction of a set of goals which, taken together, are no more important than that goal we have forsaken. If that goal is at least as important as all the others that might be competing with it, then the policy that offers us that much and more is clearly preferable to one that offers us that much or less. The upshot is that we would be obliged to strive, above all else, to avoid all-out nuclear war in the ways I have outlined.

No doubt some will deny that these circumstances do in fact obtain. But my argument requires far less than advocates of nuclear pacifism ordinarily assert and are ordinarily thought to need to assert if their case is to prevail. So even if my argument does not win over everybody, there is some reason to hope that it might attract a fair few who find the larger claims of nuclear pacifists to be just too implausible.

¹⁶ Both consequentialistic and deontological arguments arguably converge on this result (Goodin 1985a).

¹⁷ Lackey (1982, p. 205) rightly criticizes Kavka (1980) for equating unilateral nuclear disarmament with unilateral disarmament tout court. On the potential of nonnuclear defenses, see the Alternative Defence Commission (1983) report.

V

I conclude that the best thing that can be said on behalf of various pacifist proposals – unilateral nuclear disarmament and not being the first to launch a nuclear attack on the other's homeland, especially – is that they produce “moral certainties.” This does not constitute an unconditional argument for either of those policies. Both are contingent on certain empirical assumptions. Neither does this quite constitute a conclusive case for such policies, given the need to make moral trade-offs. But it is hard to imagine what trades might tempt us if “the fate of the earth” – or any appreciable chunk of it – really did hang in the balance.⁴⁰

⁴⁰ This chapter benefited from comments and criticisms of audiences at the Australian National University, the California Institute of Technology, the University of Maryland and the Aspen Conference on the Ethics of Nuclear Deterrence and Disarmament. I am particularly indebted to Brian Barry, Stanley Benn, Dave Bobrow, Arthur Burns, Bruce Cain, Frank Cowell, John Dryzek, George Eads, Steve Elkin, Jon Elster, Bob Fullinwider, Sam Gorovitz, Will Jones, Greg Kavka, Dan Kevles, Morgan Kousser, Doug Lackey, Andy Mack, Peter Menzies, Jan Narveson, Toby Page, John Passmore, Huw Price, George Quester, Richard Sylvan, Norman Schofield, Henry Shue, Jack Smart, Jack Snyder and Marc Trachtenberg.

International ethics and the environmental crisis

The problems of the environment have long been seen as global. "Only One Earth" was the motto of the 1972 Stockholm conference founding the United Nations Environmental Program and the title of the book that served as its semiofficial manifesto.¹ Even before that, attention had been firmly fixed on considerations such as the carrying capacity of the planet and the exhaustion of the earth's stock of mineral and other resources.

Still, there is something new and distinctively global about the current concern with the environment. The "first environmental crisis" was essentially a concern with problems that, though recurring the world over, could in principle be resolved perfectly well on a country-by-country basis.² When environmental problems were essentially matters of dirty air or water, they were largely matters of domestic political concern. Ill winds and shared waterways apart, pollution generally stayed in the same political jurisdiction as that in which it was generated. Of course, since all industrial nations used broadly the same dirty technologies, they all experienced similar problems of pollution. But problems that were in that sense common among a number of nations were not "shared problems" in a stronger sense, requiring concerted action among all countries for their resolution.

That is not to minimize the seriousness of the problems forming the focus of earlier environmental crusades. London's "killer fogs"

¹ Ward and Dubois 1972. See also P. Stone 1973.

² The Stockholm Declaration concedes that, inevitably, "Local and national government will bear the greatest burden for large-scale environmental policy and action within their jurisdiction," even while acknowledging a "growing class of environmental problems . . . [that] are regional or global in extent" and will "require extensive cooperation among nations and action by international organizations" (UN Conference on the Human Environment 1972, art. I, sec. 7; see similarly Ward and Dubois 1972, pts. 3 and 4, and cf. pt. 5).

were no less lethal for being purely domestic products. Neither, politically, are these traditional environmental problems necessarily all that tractable, just because they are purely domestic in nature. Even in purely domestic terms, producers with a vested interest in not cleaning up after themselves will always be a political force to be reckoned with. Still, whatever obstacles politicians face in mounting effective action against domestic polluters, those obstacles will be multiplied many times over with the addition of a genuinely international dimension to the problem.

What is striking about the environmental crisis as it is currently understood is how genuinely global it is, in contrast to traditional environmental problems. The problems at the forefront of present environmentalist discussions are problems like the degradation of the ozone layer and the "greenhouse effect." These problems are shared, internationally, in the stronger sense. They are not just problems for each nation, taken one by one. They simply cannot be resolved by isolated actions of individual nations.

London's dirty air could effectively be cleaned simply through local regulations requiring domestic users to burn smokeless coal in their fireplaces and industrial users to install scrubbers in their smokestacks. No such purely local remedies will patch the hole in the ozone layer. The voluntary decision of the U.S. government – or, indeed, the whole Organization for Economic Cooperation and Development (OECD) – to ban the use of aerosols may serve as a useful start and an important precedent; the United States produces something like 28 percent of global CFC-11 and CFC-12, and Western Europe another 30 percent all told.¹ However, if our goal is genuine stabilization of the ozone layer, and if we want to be reasonably certain of accomplishing it, then we cannot (working with present knowledge, anyway) be sufficiently sure of achieving it, even through dramatic reductions in emissions by such major producers.

In and of themselves, initiatives by single countries or even small groups of countries cannot really solve such problems. These new environmental concerns, unlike the core concerns of the "first environmental crusade," are truly global. The whole world, or some very large proportion of it, must be involved in the solution.

My argument here will build on that observation. I shall have little to say about particular environmental issues or political maneuvers surrounding them. My concern will instead be with the deeper structure of these problems, concentrating first on philosophical aspects. What alternative normative structures are logically available to us for handling such situations? Ultimately, however, this recourse to moral

¹ Wirth 1989, p. 7.

philosophy will be only incidental and instrumental. The fundamental point will be essentially political. The aim is to use philosophical insights to assist us in deciding the appropriate structure of an international regime for resolving the full range of environmental problems that we now know we face.

To foreshadow, my conclusion will be that the traditional structure of international law – guided as it is by notions of autonomous national actors with strong rights that all other national actors similarly share – is wildly inappropriate to many of these new environmental challenges. A system of shared duties or, better yet, shared responsibilities is a more fitting model, given the nature of the tasks at hand.

I

The first task, then, is to explore alternative normative structures for coping with issues of the international environment. Here I shall identify three. One is a system of shared rights, giving each nation absolute and total control over what happens within its own boundaries. Another is a system of internationally shared duties, specifying particular performances for each nation which are the duty of that nation alone; the effect is to exempt others from any obligation to pick up the slack left should any one nation fail to do its duty. A third is a system of shared responsibilities, stipulating outcomes that all nations are responsible for helping to produce; the effect there is to enjoin all nations, individually and collectively, to help take up the slack should any among them default, in whole or in part.

A. Shared rights

The fundamental principles of international law, from Grotius and Vattel forward, are all based on premises of national autonomy and noninterference with the domestic affairs of other nations. These, in turn, seem to follow from a normative structure in which each nation is thought to have a strong right to do whatever it likes to people, property and natural resources within its own jurisdiction.

Just as a system of personal rights gives individuals a "protected sphere" within which they can act without interference from others, so too does a system of international law that accords analogous rights to political entities protect the autonomy of nation-states. And just as modern liberal political theory accords maximal rights to liberty to each individual, consistent with like liberty for all, so too does liberal international law accord only such fundamental rights to any one nation as are consistent with like rights being accorded to all other

nations as well.⁴ The rights in question are, thus, shared rights – “shared” in the sense that all other agents possess rights strictly similar to one’s own.

Of course, there are limits at the margins to what liberal political theory will let agents do to themselves – as individuals or as nations. At the personal level, we standardly refuse to respect people’s decisions to sell themselves into slavery on the grounds that respecting autonomy cannot commit us to respecting decisions (even autonomously reached ones) to renounce autonomy. At the international level, we might sometimes want to impose standards of decent conduct – respecting basic human rights, for example – even upon regimes that might want autonomously to renounce them.

But those practices constitute the exceptions rather than the rule, both in liberal political theory and in the regime of international law that flows from it. By and large, if we are to interfere in the affairs of some other person or nation, we must find justification for it. That, in turn, usually amounts to showing that some of our own rights would somehow be infringed upon by the conduct in question.

If we can succeed in showing that the actions of others actually violate some of our own rights, then we can justifiably intervene in those actions, however sovereign or autonomous they may be. In the case of genuine spillovers, where others’ activities impose external costs upon us – and, crucially in this rights-based context, we actually have a right that they not impose such costs upon us – it is relatively easy to invoke notions of rights to justify our interfering with their activities. Transboundary spillovers are, within a regime of shared rights, akin to aggression, an infringement of the prerogatives of another autonomous actor with rights identical to one’s own. Thus, it is far from surprising that the case for international environmental protection long has been – and still largely continues to be – couched in terms of damage done beyond one’s own borders.⁵

Absent a demonstration of transboundary spillovers, however, we must within a regime of shared rights simply concede that environmental policy is entirely within a nation’s sovereign sphere. What is then left for us to do is to try to persuade all nations that – either because it is in their interests too, or for some other less self-serving reason – they should exercise their sovereign rights so as to produce the outcomes we want.

It is far from absurd to believe that we might be able to do so.

⁴ Hart 1955.

⁵ Note, e.g., the extent to which the “Proposed Legal Principles for Environmental Protection and Sustainable Development” adopted by the Brundtland Commission all still largely pertain to problems of “transboundary environmental interferences” (WCED 1987b).

Ward, Dubois and participants at the 1972 Stockholm Conference on the Human Environment more generally saw no real need to "reconsider national sovereignty" to solve the problems they were considering. Simply sharing information worldwide would, they supposed, be more than enough; once nations realize what environmental threats they actually face, they will have no hesitation in agreeing to concerted international action to counter them.⁶ In a similar vein, Jessica Tuchman Matthews's recent *Foreign Affairs* article attempts to cast the environmentalist case explicitly in terms of national interest, inviting nations to "redefine" conceptions of their "national security" so as to include environmental interests preeminently alongside their other "vital interests."⁷

The whole aim of rights, though, is to carve out a "protected sphere" within which agents can act with complete autonomy. What they do within that sphere – a sphere that in international law tends to be defined in basically territorial terms – is, under a regime of shared rights, purely their own business. As the much-vaunted Principle 21 of the 1972 Stockholm Declaration on the Human Environment declares, "States have, in accordance with the Charter of the United Nations and the principles of international law, the sovereign right to exploit their own resources pursuant to their own environmental policies," constrained only by the correlative "responsibility to ensure that activities within their jurisdiction . . . not cause damage to the environment of other States or . . . beyond the limits of national jurisdiction."⁸ Unless we can either show that our rights have somehow been transgressed or else persuade others to exercise their rights in line with our preferences, a regime of shared rights effectively blocks us from interfering in the actions of others – however environmentally destructive or shortsighted they might be.

B. Shared duties

Whereas classical international law revolves around notions of shared rights and sovereign prerogatives, we have recently added an overlay of shared duties. We now tend to assume, for example, that each nation must respect the fundamental human rights of its subjects, whether or not it wants to do so.⁹ Among these, it is sometimes said,

⁶ Ward and Dubos 1972, pp. 292–5. The detailed recommendations in the "Action Plan for the Human Environment" adopted at the Stockholm conference all tend toward this same implication (UN Conference on the Human Environment 1972).

⁷ Matthews 1989.

⁸ UN Conference on the Human Environment 1972.

⁹ Those duties may sometimes correlate with, or indeed derive from, the rights of

is a "fundamental right to an environment adequate for their health and well-being."¹⁰

Shared duties may correlate with – and indeed derive from – the rights of others. In the particular example just listed, they actually do so. But even where they do, they are rights of the nation's own subjects rather than of any other national actor. No other nation necessarily has any rights in the matter to press against the offending nation. That is what is crucial in differentiating this model from the last.

Under an international regime organized around notions of shared rights, the only circumstance in which there exists anything that could strictly speaking be called a "duty" to restrain any one nation's autonomous action would be if the proposed action violated the rights of some other nation. Under a regime of shared duties, by contrast, one nation may well be bound by genuine duties, even where no other nation has any strict rights.¹¹

Notice, however, that only truly fundamental duties can justifiably be imposed in this way upon nations, regardless of their particular preferences or circumstances. What follows from that fact is a rule of universality. Truly fundamental duties are equally fundamental for all agents alike. The duties thus imposed can therefore be said to be shared duties – "shared" in the sense that each nation is under the very same duties for the very same reasons as is every other nation.

For an example of this sort of normative structure drawn from the more familiar terrain of personal morality, consider the duty that each of us has to tell the truth. This is not, first and foremost, a duty that derives from any right owed to others. It rings untrue to say that the only reason we should tell the truth is simply that others have a right to be told the truth.¹² The duty is freestanding, in that sense. Furthermore, it is a duty that is imposed upon all agents alike. Whatever reasons we have for thinking that moral agents should be bound by a duty to tell the truth, those reasons are the same for all agents. The duty in question is, therefore, a shared duty.

The striking thing about duties that are shared in this way, how-

others. But those are rights of the nation's own subjects, rather than of any other national actor.

¹⁰ WCED 1987, p. 348. Principle 1 of the 1972 Stockholm Declaration puts an identical point in more florid language (UN Conference on the Human Environment 1972, pp. 1417–18).

¹¹ To sample this way of thinking, see the debate between Michael Walzer and his critics that is reprinted in Beitz, Cohen, Scanlon and Simmons (1985, pp. 165–243).

¹² Even where we think they do, it somehow rings untrue to say that the duty *derives* from the right. Surely the duty would exist, even in the absence of any particular right (or rightholder) in the matter.

ever, is that they are also very much "personalized" (or, in the current jargon, "agent-relative"). If I fail in my duty to tell you the truth, no one else is under any duty to right the wrong by disabusing you of the falsehood that I have planted in your head. The lie would be my lie; it would be charged to my moral account. Others cannot clear my account – they cannot make me any less of a liar – simply by telling the truth on my behalf. Nor can they somehow restore the moral balance of the universe by being doubly honest themselves to make up for my dishonesty.

It is indisputably true that duties such as those of truth-telling are shared duties, in the sense that everyone is under one and the same duty. The nonetheless peculiar thing about shared duties, thus construed, is that they are so weakly shared. They bind each of us individually, but none of us collectively. If one agent defaults on his duty, there is nothing any other agent should do – indeed, within this moral structure, there is nothing that any other agent even could do – to remedy the situation.

International duties are sometimes said to be like that. Consider the classic case of human rights once again. Many of those who are most anxious that their own nation respect the rights of its subjects will also insist that it would be wrong for other nations to interfere if it did not. Sometimes that position reflects simple hypocrisy, revealing that the person only halfheartedly agreed to the principle of human rights in the first place. But sometimes, at least, people urge that view because of a particular view they take about why human rights are morally important. They might think, for example, that the reason for insisting on respect for human rights has to do with a duty that nations have to display a certain attitude – an attitude of equal consideration and respect – toward their subjects. While external pressure might force a nation to perform the right actions, those would be no more than morally empty gestures if performed for the wrong reasons.¹¹ What would follow from this way of thinking about human rights is that every nation should respect the human rights of its own subjects, but that no nation should (because none usefully could) intervene if other nations failed to respect the rights of their own subjects.

That model of "tending our own garden" has been applied fairly widely by those commentators on international relations inclined to move somewhat beyond – but not too far beyond – a minimalist regime of shared rights. "Pursuing The Good" in this way, one step at

¹¹ Locke argues similarly in his *Letter Concerning Toleration* (1689) that there is no point in compelling outwardly pious religious performances from people whose hearts are not in it. Pious acts undertaken for fear of external sanction alone, unconnected to any genuine belief, will not procure a person's salvation.

a time and one country at a time, has obvious attractions. It is laudably realistic, appreciating that we cannot reasonably expect to persuade everyone in the world to do exactly the right thing at the very same instant. And if the problems in view are genuinely decomposable in that way – if they genuinely can be resolved through country-by-country action – that may well be the most effective way to pursue ‘The Good,’ in an inevitably imperfect world.

There are, however, genuine problems with that model. The first and most obvious is, of course, the simple fact that not all internationally significant problems are necessarily decomposable in that way. But that is merely an objection to the impracticality of that way of proceeding. In a more principled vein, what is perhaps morally most unattractive about this model is that it makes altogether too many concessions to realism. It lets off the hook altogether too easily those who actually do take their moral duties seriously. Under this model, they are morally in the clear just so long as they do not themselves do anything wrong. If others around them are doing wrong, even wrong of a sort that they could take action easily and costlessly to correct, they are on this account under no duty to do so. Of course, it would be good if they did. But, morally, such a performance would count as supererogatory – above and beyond the call of duty. Critics of this model might reasonably remark that, if this is so, then duty seems not to be calling loudly enough.

C. Shared responsibilities

Whereas a regime of shared duties is act-oriented, a regime of shared responsibilities is outcome-oriented.¹⁴ What duties demand of agents are specific performances. What responsibilities demand of agents are specific results, leaving the agents themselves to choose which among various possible, morally permissible actions might best achieve the mandated results. Under a system of duties, an agent is morally off the hook once he has performed precisely those actions demanded of him, even if the overall results are utterly catastrophic. A system of responsibilities does not let an agent off the hook until he has actually accomplished the prescribed ends, through some judicious choice among permissible means.

Some responsibilities – such as the responsibility of a bodyguard to protect the dignitary he or she is assigned to watch over – are peculiar to one particular agent. But many responsibilities are shared among several agents. Consider, as an example, the responsibility that

¹⁴ This model is elaborated in Chapters 5–7 in the present volume. See also Pettit and Goodin 1986.

is shared by both parents to care for their offspring. There is a single outcome in view (healthy and happy children) which the two partners, jointly and separately, are responsible for producing.

Notice, however, that the implications of sharing a responsibility are markedly different from those of sharing a duty. The principal difference derives, in turn, from the difference between act-oriented and result-oriented moral systems. To share a duty is to have a duty just like another's; but it is still very much your own duty, and if you fail to do it, no one else can do it for you. To share a responsibility, by contrast, is to be responsible together with various others for producing certain outcomes. And since it is the outcome that each is responsible for producing, there is usually something each can (and should, if possible) do to pick up the slack, if any of the others default in their responsibility. Thus, for example, parents are, insofar as they are able, jointly and separately responsible for ensuring that their children's basic needs are met. What that means, in turn, is that each parent is responsible for assuming complete responsibility for catering to the basic needs of the couple's children, should the other partner prove unable or unwilling to shoulder his or her share of the burden.

That leads to the second important point of difference with a regime of shared responsibilities. Under the other sorts of regimes, it is inappropriate (wrong, under a regime of shared rights; pointless, under a regime of shared duties) to force people to do what, in some larger sense, they should. Under a regime of shared responsibilities, by contrast, it is perfectly proper to do what one can – within limits – to force others to shoulder their share of a responsibility that is jointly shared. It is your business to do so, precisely because their default would increase the share of the burden that would morally fall to you to bear. Thus, in the example of family relations, it is thought to be perfectly proper to use the force of the law to extract child support payments from financially solvent parents who have chosen to leave their families.

Transpose this family model, now, from hearthside settings to the international arena. What would it mean to say, then, that what nations shared were genuine responsibilities rather than mere duties? First of all, it would fix the moral focus on the outcomes that they were collectively supposed to produce, rather than on specific acts of specific agents. Second, it would mean that each nation would be responsible for making good any shortcomings, should other nations fail to do their full part toward producing those ends. And third, it would mean that each could properly press others to do their part toward producing those shared ends.

In the context of international human rights policy, for example, a regime of shared responsibilities would have quite clear and distinc-

tive implications. Under such a regime, it most definitely is the business of the international community as a whole to ensure that states respect human rights, even if they are not so inclined. It would be legitimate for other nations to do whatever they can, within limits, to force delinquent nations to respect human rights. And it would be not only morally permissible but morally mandatory for nations, insofar as possible, to respect human rights on behalf of any delinquent nation – by offering political asylum to that nation's persecuted subjects, for example.

In the context of international environmental policy, a regime of shared responsibilities would imply, first of all, that it is morally permissible for environmentally conscientious nations to bring pressure, at least in certain ways, upon nations that fail to discharge their environmental responsibilities. It would be perfectly permissible, rather than a gross infringement of another nation's sovereign prerogatives, for one nation to grant licenses to fish in its territorial waters only to the ships of nations that comply with international standards to protect fish stocks within their own territorial waters. It would also be perfectly permissible for bilateral or multilateral aid donors to attach strings to loans, making receipt conditional upon effective policies to protect the environment within the recipient nations.

Such a model would imply, secondly, that it would be both fitting and proper for environmentally conscientious nations to do double-duty, should others refuse to do their duty at all. If some nations are not going to do their part, then the others must do more than their share if the task is going to get done at all. On this analysis, therefore, there should be no moral qualms about paying Brazil to stop destroying the Amazon rain forests, even though this would amount to paying Brazilians to do no more than what morally they should be doing anyway. And on this analysis, it is perfectly proper for environmentally conscientious nations to overcomply with international agreements protecting the environment – reducing their whale catch or their production of chlorofluorocarbons (CFCs) or "greenhouse gases" by more than the treaty requires – once it becomes clear that some other nations are going to undercomply.

D. Mixed models

Naturally, these are all highly stylized models and the distinctions between them tend to blur in practice. It is nevertheless worth setting out distinctions as clearly as possible, even at the risk of some artificiality, so that the advantages and disadvantages of any particular component in the larger mixture can be clearly assessed.

While conceding that actual cases may always be mixed, it would

be a mistake, however, to jump to the conclusion that actual cases always will inevitably be mixed. It is commonly said, for example, that rights entail responsibilities; and that might lead us to suppose that those two models of international ethics are necessarily complementary rather than competing ones. However, whether or not that is true depends on what account is given of the entailment relationship. According to one very standard interpretation, rights entail just a responsibility to respect the analogous rights of others. If that is all there is to the relationship, then the rights and corollary responsibilities both work strictly within one and the same model of shared rights, as described above in Section IA. There are, of course, other ways of interpreting the rights-responsibilities entailment relationship. One, for example, deals in terms of the duties that the powerful have to protect the less powerful.¹⁵ But genuinely distinctive responsibilities, akin to those to be imposed under the model of shared responsibilities described in Section IC, will arise only on some such stronger and more contentious accounts of the entailment relationship.

II

With all this philosophical apparatus in place, let us return to the actual policy problems – problems of the environment – that motivated this inquiry in the first place. Of course, there are many problems with the environment, each subtly different from the other. Different sorts of policy responses, and different structures of international regime, are therefore going to be best suited to solving all the various problems of the global environment.

Let us, however, try to cut through all those subtleties and focus instead on fundamentals. Different as they may be in other respects, notions of shared rights and notions of shared duties both deal in terms of the actions of nations one at a time. Some, perhaps many, environmental problems are indeed decomposable in that way. The more nations there are implementing a policy, the more likely it is that the desired outcome will be achieved; and the relationship is thus a smoothly increasing function of how many, and to what extent, actors are working toward that end. In such a case, isolated actions of individual nations are, in principle, perfectly capable of producing – or at least of contributing usefully to the production of – the desired outcome. And it is therefore perfectly defensible for us to pursue those goals through normative structures focusing on the actions of nations one at a time.

Some of the most worrisome environmental problems are not like

¹⁵ Goodin 1985c.

that at all, however. Instead, they are more akin to "lumpy public goods." Instead of policy inputs translating smoothly into environmental outputs, the response curve is more of a "step function"; inputs must pass a certain threshold before they make any difference whatsoever to the outcome.¹⁶ On one plausible account alluded to earlier, ozone depletion and resulting climate change might be like that.

In such cases, concerted action among a large group of countries will be required to make any difference at all to the outcome, and normative structures focusing on the isolated acts of single states are wildly inappropriate to the situation. For these second-wave problems that characterize the "new environmental crisis," therefore, a regime of shared responsibilities is the normative structure that is prudentially required.

III

To say that a regime of shared responsibilities is the normative structure that the situation requires, however, is not to say that it will be set in place automatically. Politically, we must start from where we find ourselves, and that is in a world of sovereign states. Even the most committed environmentalist must take due account of that fact.¹⁷ In such a setting, shared responsibilities can acquire practical political force only if (and only to the extent that) they are recognized by nations themselves, through treaties and other similar international instruments.

There are good grounds for suspecting that this strategy is morally suboptimal, second-best, or worse. Of course, it is perfectly possible in a regime of shared rights for states, through the exercise of their sovereign prerogatives, to sign treaties assuming various responsibilities to be shared with other cosignatories. But the shared responsibilities that emerge in that way are very different from those involved in models built around those notions directly. Whereas the shared responsibilities under those latter models would be foundational, treaty-based responsibilities would instead be merely derivative – de-

¹⁶ For an application of such a model to environmental problems, see M. Taylor and Ward 1982.

¹⁷ As does the Stockholm Declaration when saying, "International matters concerning the protection . . . of the environment should be handled in a cooperative spirit by all countries, big or small, on an equal footing. Cooperation through multilateral or bilateral arrangements or other appropriate means is essential," but it must work "in such a way that due account is taken of the sovereignty and interests of all states" (UN Conference on the Human Environment 1972, Principle 24).

ivative from the rights that sovereign states have to sign such agreements with other sovereign states.

The disadvantage of their being derivative, in turn, is that they are virtually always revocable, at least in principle. What sovereign states do through the exercise of their sovereign prerogatives they can typically undo in the same manner. As the old saying in constitutional law has it, sovereigns cannot bind their future selves – or at least they cannot do so without undermining the sovereignty of those future sovereigns.¹⁸

Under regimes based directly upon notions of shared responsibilities, by contrast, the responsibilities are regarded instead as foundational. Having in that way an existence independent of the actions of sovereign states, they cannot simply be revoked at the pleasure of the states concerned in quite the same way that responsibilities deriving merely from treaties typically can. Such is the great disadvantage of deriving shared responsibilities from treaty commitments alone.

Even if that treaty-based strategy is morally suboptimal, though, at least it has realism to recommend it. Given where we are starting – in a world of sovereign states – perhaps the treaty-based strategy is the only way to move toward a regime of shared responsibilities.

Still, if a regime of shared responsibilities can only emerge in present circumstances from treaties, not all sorts of treaties are equally good for the purpose. Some treaties institute a regime of shared responsibilities, whereas others just serve to reinforce regimes of shared rights.

Compare, for example, the 1985 Vienna Convention for the Protection of the Ozone Layer with the 1987 Montreal Protocol on Substances that Deplete the Ozone Layer.¹⁹ Notice how the Montreal Protocol was set to go into force, not (as with the Vienna Convention) when a fixed number of signatories ratified it, but rather when it was ratified by countries accounting for two-thirds of the estimated 1986 consumption of ozone-depleting substances. The idea – which, in terms of a model of shared responsibilities, is obviously the right idea – is that what matters is not how many members there are in the club but whether the members that are in it have the capacity to make the relevant difference to the outcomes.²⁰ Or notice, again, how rather than just encouraging systematic observation, research and informa-

¹⁸ Sometimes of course treaties do explicitly renounce sovereignty in certain respects, in which case the obligations arising under them may well be irrevocable.

¹⁹ Vienna Convention 1985. Montreal Protocol 1987. For the purely illustrative purposes here, I simply gloss over the fact that the latter is a Protocol concluded under the former Convention; the differences here described may reflect no more than the inevitably different levels of generality in such different documents.

²⁰ Kennan 1970.

tion exchange, as under the Vienna Convention, the Montreal Protocol actually imposes some rather onerous burdens upon signatory states, committing them first to freezing and then to reducing sharply their emissions of ozone-depleting substances. Or notice, yet again, how the Montreal Protocol commits signatories, in a way the Vienna Convention patently does not, to attempting to influence the ozone-depleting actions of other, nonsignatory states. That might be regarded by advocates of a shared-rights model of international relations as an unwarranted interference with the domestic affairs of another nation, but it would be perfectly permissible in pursuit of genuinely shared moral responsibilities. In all these respects, the Montreal Protocol is a model of how to – and the Vienna Convention a model of how not to – draft treaties institutionalizing a regime of shared responsibilities.”

The primary recommendation of this chapter – which for reasons given at the outset of this section is still very much a second-best solution, morally – is for the recognition of such responsibilities through many more treaties along the lines of an extended version of the Montreal Protocol. Pending such international developments, however, there are still useful steps that individual nations can take. A model of shared moral responsibilities for environmental protection would, for example, legitimize a nation refusing to allow the manufacture or export of CFCs or the technology to produce them. Or, for another example, it would legitimize a nation unilaterally refusing to provide aid or loans to countries that manufacture CFCs.

Other nations may protest that this constitutes interference in their own domestic affairs. And of course in a way it does. But that objection bites only if we are thinking in terms of rights of sovereign states. The point of this chapter is that this is the wrong way to be thinking about the new wave of environmental concerns.”

“ In other respects, however, the Montreal Protocol is less than a perfect paradigm. Instead of requiring fixed performances from each nation, a regime of genuinely shared responsibilities should stipulate that the more signatories there are and the more emissions they account for, the less each should have to pay. (Obligations under that model would be more like those falling to “names” in the Lloyds insurance syndicates.) The Montreal Protocol does, however, provide for regular meetings of signatories to update the list of ozone-depleting substances and requirements for their control, in light of subsequent research. Perhaps that is in practice the mechanism by which such adjustments would best be made in any case.

“ I am grateful to James Crawford, Michael Glendon and Joel Rosenthal and participants at a symposium at the University of California, Davis, for comments on an earlier draft of this chapter.

References

- Abel, Richard. 1982. Torts. Pp. 185-200 in David Kairys, ed., *The politics of law*. New York: Pantheon.
- Abel-Smith, Brian, and Kay Titmuss, eds. 1987. *The philosophy of welfare: selected writings of Richard M. Titmuss*. London: Allen & Unwin.
- Ackerman, Bruce A. 1977. *Private property and the constitution*. New Haven, Conn.: Yale University Press.
- Adams, Robert Merrihew. 1976. Motive utilitarianism. *Journal of Philosophy* 73: 467-81.
- Allais, Maurice, and Ole Hagen, eds. 1979. *Expected utility hypothesis and the Allais paradox*. Dordrecht: Reidel.
- Alternative Defence Commission. 1983. *Defence without the bomb*. London: Taylor & Francis.
- American Law Institute. 1986. Model eminent domain code. Vol. 13 of *Uniform laws annotated*. St. Paul, Minn.: West, for ALI.
- Anonymous. 1983. Developments in the law: immigration policy and the rights of aliens. *Harvard Law Review* 96: 1286-1465.
- Anscombe, G. E. M. 1958. Modern moral philosophy. *Philosophy* 33: 1-19.
- Anscombe, G. E. M. 1978. On the source of the authority of the state. *Ratio* 20: 1-28.
- Arneson, Richard J. 1980. Mill versus paternalism. *Ethics* 90: 470-89.
- Arrow, Kenneth J. 1963. *Social choice and individual values*. 2nd ed. New Haven, Conn.: Yale University Press; originally published 1951.
- Arrow, Kenneth J. 1963. Uncertainty and the welfare economics of medical care. *American Economic Review* 53: 941-73.
- Arrow, Kenneth J. 1967. The place of moral obligation in preference systems. Pp. 117-19 in Sidney Hook, ed., *Human values and economic policy*. New York: New York University Press.
- Arrow, Kenneth J., and Anthony C. Fisher. 1974. Environmental preservation, uncertainty and irreversibility. *Quarterly Journal of Economics* 88: 312-19.
- Atiyah, P. S. 1975. *Accidents, compensation and the law*. 2nd ed. London: Weidenfeld & Nicolson.
- Atiyah, P. S. 1980. *Accidents, compensation and the law*. 3rd ed. London: Weidenfeld & Nicolson.

References

- Atiyah, P. S., and Peter Cane. 1987. *Atiyah's accidents, compensation and the law*. 4th ed. London: Weidenfeld & Nicolson.
- Atkinson, A. B. 1975. *The economics of inequality*. Oxford: Clarendon Press.
- Atkinson, A. B. 1990. Income maintenance for the unemployed in Britain and the response to high unemployment. *Ethics* 100: 569-85.
- Austin, J. L. 1956-7. A plea for excuses. *Proceedings of the Aristotelian Society* 57: 1-30.
- Avineri, Shlomo, and Avner de-Shalit, eds. 1992. *Communitarianism and individualism*. Oxford: Oxford University Press.
- Baier, Kurt. 1970. Responsibility and action. Pp. 100-16 in Myles Brand, ed., *The nature of human action*. Glenview, Ill.: Scott Foresman.
- Baker, John. 1992. An egalitarian case for basic income. In Van Parijs 1992, pp. 101-27.
- Bale, Malcolm D., and John H. Mutti. 1978. Income losses, compensation and international trade. *Journal of Human Resources* 13: 278-85.
- Ball, Desmond. 1981. *Can nuclear war be controlled?* Adelphi papers no. 169. London: International Institute for Strategic Studies.
- Banfield, Edward, and James Q. Wilson. 1964. Public regardingness as a value premise in voting behavior. *American Political Science Review* 63: 876-87.
- Barr, Nicholas. 1987. *The economics of the welfare state*. London: Weidenfeld & Nicolson.
- Barr, Nicholas. 1989. Social insurance as an efficiency device. *Journal of Public Policy* 9: 59-82.
- Barry, Brian. 1964. The public interest. *Proceedings of the Aristotelian Society (Supplement)* 38: 1-18.
- Barry, Brian. 1965. *Political argument*. London: Routledge & Kegan Paul.
- Barry, Brian. 1975. The consociational model and its dangers. *European Journal of Political Research* 3: 393-412. Reprinted in Barry 1989a, 136-55.
- Barry, Brian. 1979a. And who is my neighbor? *Yale Law Journal* 88:629-58. Reprinted in Barry 1989a, 322-59.
- Barry, Brian. 1979b. Justice as reciprocity. Pp. 50-78 in Eugene Kamenka and Alice E.-S. Tay, eds., *Justice*. London: Edward Arnold. Reprinted in Barry 1989a, pp. 463-93.
- Barry, Brian. 1982. Humanity and justice in global perspective. Pp. 219-52 in J. R. Pennock and J. W. Chapman, eds., *Nomos XXIV: ethics, economics and the law*. New York: New York University Press. Reprinted in Barry 1989a, pp. 434-62.
- Barry, Brian. 1983. Self-government revisited. Pp. 121-54 in David Miller and Larry Siedentop, eds., *The nature of political theory*. Oxford: Clarendon Press. Reprinted in Barry 1989a, pp. 156-86.
- Barry, Brian. 1986. Lady Chatterley's lover and Doctor Fischer's bomb party: liberalism, Pareto optimality and the problem of objectionable preferences. In Elster and Hylland 1986, pp. 11-43. Reprinted in Barry 1989a, pp. 360-91.
- Barry, Brian. 1989a. *Democracy, power and justice*. Oxford: Clarendon Press.
- Barry, Brian. 1989b. Utilitarianism and preference change. *Utilitas* 1: 278-82.
- Bator, F. M. 1957. The simple analytics of welfare maximization. *American Economic Review* 47: 22-59.

References

- Beale, J. H., Jr. 1895. Recovery for the consequences of an act. *Harvard Law Review* 9: 80-9.
- Beitz, Charles. 1979. *Political theory and international relations*. Princeton, N.J.: Princeton University Press.
- Beitz, Charles. 1983. Cosmopolitan ideals and national sentiment. *Journal of Philosophy* 80: 591-600.
- Beitz, Charles; Marshall Cohen; Thomas Scanlon; and A. John Simmons, eds. 1985. *International ethics*. Princeton, N.J.: Princeton University Press.
- Benn, Stanley I. 1971. Privacy, freedom and respect for persons. Pp. 1-26 in J. R. Pennock and J. W. Chapman, eds., *Nomos XII: privacy*. New York: Atherton.
- Benn, Stanley I. 1976. Rationality and political behaviour. Pp. 246-67 in G. W. Mortimore and S. I. Benn, eds., *Rationality and the social sciences*. London: Routledge & Kegan Paul.
- Benn, Stanley I. 1979. The problematic rationality of political participation. Pp. 291-312 in Peter Laslett and James S. Fishkin, eds., *Philosophy, politics and society*. 5th series. Oxford: Blackwell.
- Benn, Stanley I. 1983. Deterrence or appeasement? *Journal of Applied Philosophy* 1: 5-19.
- Benn, Stanley I. 1988. *A theory of freedom*. Cambridge: Cambridge University Press.
- Bennett, Jonathan. 1983. Positive and negative relevance. *American Philosophical Quarterly* 20: 185-94.
- Bennett, Jonathan. 1989. Two departures from consequentialism. *Ethics* 100: 54-66.
- Bennett, Robert. 1979. "Mere" rationality in constitutional law: judicial review and democratic theory. *California Law Review* 67: 1049-1103.
- Bentham, Jeremy. 1843. *Works*, ed. John Bowring. Edinburgh: William Tait.
- Bentham, Jeremy. 1970. *An introduction to the principles of morals and legislation*, ed. J. H. Burns and H. L. A. Hart. London: Athlone Press; originally published 1789.
- Bentham, Jeremy. 1988. *A fragment on government*, ed. J. H. Burns and H. L. A. Hart. Cambridge: Cambridge University Press; originally published 1776.
- Berger, Peter L., and Richard J. Neuhaus. 1977. *To empower people*. Washington, D.C.: American Enterprise Institute.
- Beveridge, William H. 1942. *Social insurance and allied services*. Cmd. 6404. London: HMSO.
- Bickel, Alexander M. 1975. *The morality of consent*. New Haven, Conn.: Yale University Press.
- Bilmes, Jack M. 1979. The evolution of decisions in a Thai village: a quasi-experimental study. *Human Organization* 38: 169-78.
- Black, Hugo. 1960. Majority opinion. *Armstrong v. U.S.* 364 U.S. 40.
- Blaug, Marc. 1963. The myth of the old poor law and the making of the new. *Journal of Economic History* 23: 151-84.
- Blechman, Barrie M., and Douglas M. Hart. 1982. The political utility of nuclear weapons: the 1973 Middle East crisis. *International Security* 7: 132-56.
- Blough, Thomas. 1941. The role of compensation in the war-economic system. *Bulletin of the Oxford Institute of Statistics* 3: 99-103.

References

- Blumer, Martin. 1987. *The social basis of community care*. London: Unwin Hyman.
- Bracken, Paul. 1983. *The command and control of nuclear forces*. New Haven, Conn.: Yale University Press.
- Braithwaite, V. A. 1990. *Bound to care*. Sydney: Allen & Unwin.
- Brandt, Richard B. 1959. *Ethical theory*. Englewood Cliffs, N.J.: Prentice-Hall.
- Brandt, Richard B. 1963. Towards a credible form of utilitarianism. Pp. 107–43 in Hector-Neri Castañeda and George Nakhnikian, eds., *Morality and the language of conduct*. Detroit: Wayne State University Press.
- Brandt, Richard B. 1967. Personal values and the justification of institutions. Pp. 22–40 in Stanley Hook, ed., *Human values and economic policy*. New York: New York University Press.
- Brandt, Richard B. 1972. Utilitarianism and the rules of war. *Philosophy and Public Affairs* 1: 145–65. Reprinted in Cohen et al. 1974, pp. 25–45, and Brandt 1992, pp. 336–53.
- Brandt, Richard B. 1979. *A theory of the good and the right*. Oxford: Clarendon Press.
- Brandt, Richard B. 1985. A motivational theory of excuses in the criminal law. Pp. 165–200 in J. R. Pennock and J. W. Chapman, eds., *Nomos XXVII: criminal justice*. New York: New York University Press. Reprinted in Brandt 1992, pp. 235–62.
- Brandt, Richard B. 1988. Fairness to indirect optimific theories in ethics. *Ethics* 98: 341–61. Reprinted in Brandt 1992, pp. 137–57.
- Brandt, Richard B. 1992. *Morality, utilitarianism and rights*. Cambridge: Cambridge University Press.
- Brandt, Willy. 1980. *North-south: a programme for survival*. London: Pan.
- Braybrooke, David. 1987. *Meeting needs*. Princeton, N.J.: Princeton University Press.
- Braybrooke, David, and C. E. Lindblom. 1963. *A strategy of decision*. New York: Free Press.
- Brennan, Geoffrey. 1973. Pareto-optimal redistribution: the case of malice and envy. *Journal of Public Economics* 2: 173–83.
- Brennan, Geoffrey, and Loren Lomasky. 1993. *Democracy and decision*. Cambridge: Cambridge University Press.
- Brewer, Daniel. 1893. Majority opinion. *Monongahela Navigation Co. v. U.S.* 148 U.S. 312–44.
- Brierly, J. L. 1936. *The law of nations*. 2nd ed. Oxford: Clarendon Press.
- Brittan, Samuel. 1988. *A restatement of economic liberalism*. 2nd ed. London: Macmillan.
- Broad, C. D. 1916. On the function of false hypotheses in ethics. *International Journal of Ethics* (now *Ethics*) 26: 377–97.
- Brody, Richard A. 1968. Deterrence. In Sills 1968, vol. 4, pp. 130–3.
- Broome, John. 1978a. Choice and value in economics. *Oxford Economic Papers* 30: 313–33.
- Broome, John. 1978b. Trying to value a life. *Journal of Public Economics* 9: 91–100.
- Browning, E. K. 1975. The externality argument for in-kind transfers: some critical remarks. *Kyklos* 28: 526–44.

References

- Brownlie, Ian. 1966. *Principles of public international law*. Oxford: Clarendon Press.
- Buen, E. H. 1982. *Modern law of real property*. London: Butterworth.
- Burns, Arthur Lee. 1970. *Ethics and deterrence*. Adelphi papers no. 96. London: International Institute for Strategic Studies.
- Burrows, William E. 1984. Ballistic missile defense: the illusion of security. *Foreign Affairs* 62: 843-56.
- Calabresi, Guido. 1970. *The cost of accidents*. New Haven, Conn.: Yale University Press.
- Calabresi, Guido, and A. Douglas Melamed. 1972. Property rules, liability rules and inalienability: one view of the cathedral. *Harvard Law Review* 85: 1089-1128.
- Campbell, Angus. 1972. Aspiration, satisfaction and fulfillment. Pp. 441-66 in A. Campbell and Philip E. Converse, eds., *The human meaning of social change*. New York: Russell Sage Foundation.
- Cannon, C. 1963. *Procedure in the House of Representatives*. 87th Congress, 2nd Session. Washington, D.C.: Government Printing Office.
- Cardozo, Benjamin. 1915. Majority opinion. *People v. Crane*. 214 N.Y. 154; 108 N.E. 427-33.
- Castles, Francis G. 1985. *The working class and the welfare state*. Sydney: Allen & Unwin.
- Chambers, David L. 1979. *Making fathers pay: enforcement of child support*. Chicago: University of Chicago Press.
- Chase, Harold W., and Craig R. Ducat. 1980. *1980 supplement to Edward S. Corwin's The Constitution, and what it means today*. Princeton, N.J.: Princeton University Press.
- Cohen, Joshua. 1989. Deliberation and democratic legitimacy. Pp. 17-34 in Alan Hamlin and Philip Pettit, eds., *The good polity*. Oxford: Blackwell.
- Cohen, Marshall, Thomas Nagel, and Thomas Scanlon, eds. 1974. *War and moral responsibility*. Princeton, N.J.: Princeton University Press.
- Cook, Philip J., and Daniel A. Graham. 1977. The demand for insurance and protection: the case of irreplaceable commodities. *Quarterly Journal of Economics* 91: 143-56.
- Cooter, Robert, and Peter Rappoport. 1984. Were the ordinalists wrong about welfare economics? *Journal of Economic Literature* 22: 507-30.
- Cordes, Joseph J. 1979. Compensation through relocation assistance. *Land Economics* 55: 486-98.
- Cordes, Joseph J., and Robert S. Goldfarb. 1983. Alternative rationales for severance pay compensation under airline deregulation. *Public Choice* 41: 351-69.
- Corwin, Edward S. 1911. The doctrine of due process of law before the Civil War. *Harvard Law Review* 24: 366-85, 460-79.
- Corwin, Edward S. 1978. *The Constitution, and what it means today*, ed. Harold W. Chase and Craig R. Ducat. 14th ed. Princeton, N.J.: Princeton University Press.
- Cottingham, John. 1986. Partiality, favouritism and morality. *Philosophical Quarterly* 36: 357-73.
- Culyer, A. J. 1974. Economics, social policy and disability. Pp. 17-29 in Dennis

References

- Lees and Stella Shaw, eds., *Impairment, disability and handicap*. London: Heinemann for the SSRC.
- d'Abbs, Peter. 1982. *Social support networks: a critical review of models and findings*. Melbourne: Institute of Family Studies.
- Daniels, Norman. 1985. *Just health care*. Cambridge: Cambridge University Press.
- Danziger, Sheldon, Robert Haveman, and Robert Plotnick. 1981. How income transfer programs affect work, savings and income distribution. *Journal of Economic Literature* 19: 975-1028.
- Dasgupta, Partha. 1982. Utilitarianism, information and rights. In Sen and Williams 1982, pp. 199-218.
- Dasgupta, Partha, and Debraj Ray. 1986-7. Inequality as a determinant of malnutrition and unemployment. *Economic Journal* 96: 1011-34 and 97: 177-88.
- Davidson, Donald. 1980. *Essays on actions and events*. Oxford: Clarendon Press.
- Davidson, Donald. 1986. Judging interpersonal interests. In Elster and Hylland 1986, pp. 195-212.
- Davidson, Donald. 1987. Knowing one's own mind. *Proceedings and Addresses of the American Philosophical Association* 60: 4411-58.
- Davidson, Donald, and Patrick Suppes. 1957. *Decision making*. Stanford, Calif.: Stanford University Press.
- Davis, Nancy (Ann). 1980a. The priority of avoiding harm. Pp. 172-214 in Bonnie Steinbock, ed., *Killing and letting die*. Englewood Cliffs, N.J.: Prentice-Hall.
- Davis, Nancy (Ann). 1980b. Utilitarianism and responsibility. *Ratio* 22: 15-35.
- Day, J. P. 1981. Compensatory discrimination. *Philosophy* 56: 55-72.
- Deacon, Alan, and Jonathan Bradshaw. 1983. *Reserved for the poor: the means test in British social policy*. Oxford: Blackwell and Martin Robertson.
- Deigh, John. 1984. On the right to be punished: some doubts. *Ethics* 94: 191-211.
- Dennett, Daniel, and Karel Lambert, eds. 1978. *The philosophical lexicon*. 7th ed. Mimeograph copy, Department of Philosophy, Tufts University.
- Derthick, Martha. 1979. *Policymaking for social security*. Washington, D.C.: Brookings Institution.
- Dickens, Charles. 1969. *Hard times*. Harmondsworth: Penguin; originally published 1854.
- Donagan, Alan. 1968. Is there a credible form of utilitarianism? Pp. 187-202 in Michael Bayles, ed., *Contemporary utilitarianism*. Garden City, N.Y.: Anchor Doubleday.
- Donagan, Alan. 1977. *The theory of morality*. Chicago: University of Chicago Press.
- Douglas, Mary, and Brian Isherwood. 1979. *The world of goods*. London: Allen Lane.
- Downs, Anthony. 1957. *An economic theory of democracy*. New York: Harper.
- Doyal, Len, and Ian Gough. 1984. A theory of human needs. *Critical Social Policy* 4(1): 6-38.
- Drèze, Jean, and Amartya Sen. 1989. *Hunger and public action*. Oxford: Clarendon Press.

- Dryzek, John. 1983. Present choices, future consequences. *World Futures* 19: 1-19.
- Dryzek, John. 1990. *Discursive democracy*. Cambridge: Cambridge University Press.
- Duncan, Greg J., with Richard D. Coe, Mary E. Corcoran, Martha S. Hill, Saul D. Hoffman, and James N. Morgan. 1984. *Years of poverty, years of plenty*. Ann Arbor: Institute for Social Research, University of Michigan.
- Duncan, Otis Dudley. 1975. Does money buy satisfaction? *Social Indicators Research* 2: 267-74.
- Dunleavy, Patrick, and Hugh Ward. 1981. Exogenous voter preferences and parties with state power. *British Journal of Political Science* 11: 351-80.
- Dworkin, Gerald. 1971. Paternalism. Pp. 107-26 in Richard A. Wasserstrom, ed., *Morality and the law*. Belmont, Calif.: Wadsworth.
- Dworkin, Gerald. 1982. Is more choice better than less? *Midwest Studies in Philosophy* 7: 47-61. Reprinted in Dworkin 1988, pp. 62-84.
- Dworkin, Gerald. 1985. Nuclear intentions. *Ethics* 95: 445-60. Reprinted in Hardin et al. 1985, pp. 37-52.
- Dworkin, Gerald. 1988. *The theory and practice of autonomy*. Cambridge: Cambridge University Press.
- Dworkin, Ronald M. 1977. *Taking rights seriously*. London: Duckworth.
- Dworkin, Ronald M. 1978. Liberalism. Pp. 113-43 in Stuart Hampshire, ed., *Public and private morality*. Cambridge: Cambridge University Press. Reprinted in Dworkin 1985, pp. 181-204.
- Dworkin, Ronald M. 1980. Why efficiency? *Hofstra Law Review* 8: 563-90. Reprinted in Dworkin 1985, pp. 267-89.
- Dworkin, Ronald M. 1981. What is equality? *Philosophy and Public Affairs* 10: 185-246, 283-345.
- Dworkin, Ronald M. 1985. *A matter of principle*. Cambridge, Mass.: Harvard University Press.
- Eckstein, Otto. 1961. A survey of the theory of public expenditure criteria. Pp. 439-94 in National Bureau of Economic Research, ed., *Public finances: needs, sources and utilization*. Princeton, N.J.: Princeton University Press.
- Edwards, Meredith. 1981. *Financial arrangements within families*. Canberra: National Women's Advisory Council.
- Edwards, Paul, ed. 1967. *Encyclopedia of philosophy*. 8 vols. New York: Macmillan.
- Elias, Robert. 1983. *Victims of the system: crime victims and compensation in American politics and criminal justice*. New Brunswick, N.J.: Transaction Books.
- Elliot, Robert. 1982. Faking nature. *Inquiry* 25: 81-94.
- Ellis, Adrian, and Anthony Heath. 1983. Positional competition, or an offer you can't refuse? Pp. 1-22 in A. Ellis and Krishan Kumar, eds *Dilemmas of liberal democracies*. London: Tavistock.
- Ellsberg, Daniel. 1961. Risk, ambiguity and the Savage axioms. *Quarterly Journal of Economics* 75: 643-69.
- Ellsberg, Daniel. 1968. *The theory and practice of blackmail*. Rand paper P-3883. Santa Monica, Calif.: Rand Corporation.
- Elster, Jon. 1976a. Boudon, education and the theory of games. *Social Science Information* 15: 733-40.

References

- Elster, Jon. 1976b. Some conceptual problems in political theory. Pp. 245-70 in Brian Barry, ed., *Power and political theory*. London: Wiley.
- Elster, Jon. 1978. *Logic and society*. London: Wiley.
- Elster, Jon. 1979a. Risk, uncertainty and nuclear power. *Social Science Information* 18: 371-400.
- Elster, Jon. 1979b. *Ulysses and the sirens*. Cambridge: Cambridge University Press.
- Elster, Jon. 1982. Utilitarianism and the genesis of wants. In Sen and Williams 1982, pp. 219-38. Reprinted in Elster 1983, pp. 109-40.
- Elster, Jon. 1983. *Sour grapes*. Cambridge: Cambridge University Press.
- Elster, Jon, and Aanund Hylland, eds. 1986. *Foundations of social choice theory*. Cambridge: Cambridge University Press.
- Ely, John Hart. 1970. Legislative and administrative motivation in constitutional law. *Yale Law Journal* 79: 1205-1314.
- Enker, Arnold N. 1974. A comparative review of compensation for victims of crime. Vol. 2, pp. 121-38 in Israel Drapkin and Emilio Viano, eds., *Victimology: a new focus*. Lexington, Mass.: Lexington Books, D. C. Heath.
- Epstein, Richard A. 1985. *Takings: private property and the power of eminent domain*. Cambridge, Mass.: Harvard University Press.
- Esping-Andersen, Gøsta. 1990. *The three worlds of welfare capitalism*. Princeton, N.J.: Princeton University Press.
- Etzioni, Amitai. 1988. *The moral dimension*. New York: Free Press.
- Feinberg, Joel. 1966. Duties, rights and claims. *American Philosophical Quarterly* 3: 137-44. Reprinted in Feinberg 1980, pp. 130-42.
- Feinberg, Joel. 1970. *Doing and deserving*. Princeton, N.J.: Princeton University Press.
- Feinberg, Joel. 1973. *Social philosophy*. Englewood Cliffs, N.J.: Prentice-Hall.
- Feinberg, Joel. 1974. Noncomparative justice. *Philosophical Review* 83: 297-338. Reprinted in Feinberg 1980, pp. 265-306.
- Feinberg, Joel. 1980. *Rights, justice and the bounds of liberty*. Princeton, N.J.: Princeton University Press.
- Feinberg, Joel. 1986. *Harm to self*. New York: Oxford University Press.
- Feldstein, Martin. 1976a. Compensating in tax reform. *National Tax Journal* 29: 123-30.
- Feldstein, Martin. 1976b. On the theory of tax reform. *Journal of Public Economics* 6: 77-104.
- Finnis, John, Joseph M. Boyle, Jr., and Germain Grisez. 1987. *Nuclear deterrence, morality and realism*. Oxford: Clarendon Press.
- Fiorina, Morris P. 1981. *Retrospective voting in American national elections*. New Haven, Conn.: Yale University Press.
- Fischhoff, Baruch. 1983. Strategic policy preferences. *Journal of Social Issues* 39: 133-60.
- Fisher, Adrian S., chief reporter. 1965. *Restatement (second) of the foreign relations law of the United States*. St. Paul, Minn.: American Law Institute.
- Fishkin, James S. 1982. *The limits of obligation*. New Haven, Conn.: Yale University Press.
- Fletcher, George P. 1968. Legal aspects of the decision not to prolong life. *Journal of the American Medical Association* 203: 65-8.

References

- Fletcher, George P. 1979. Legality as privacy. Pp. 182-207 in Robert L. Cunningham, ed., *Liberty and the rule of law*. College Station: Texas A & M Press.
- Fletcher, George P. 1993. *Loyalty: an essay on the morality of relationships*. New York: Oxford University Press.
- Føllesdal, Dagfin. 1979. Some ethical aspects of recombinant DNA research. *Social Science Information* 18: 401-19.
- Frankena, William K. 1977. Moral philosophy and world hunger. Pp. 66-84 in William Aitken and Hugh La Follette, eds., *World hunger and moral obligation*. Englewood Cliffs, N.J.: Prentice-Hall.
- Frankfurt, Harry G. 1971. Freedom of the will and the concept of a person. *Journal of Philosophy* 68: 5-20.
- Frankfurt, Harry G. 1984. Necessity and desire. *Philosophy and Phenomenological Research* 14: 1-13.
- Freedman, Michael. 1978. *The new liberalism*. Oxford: Clarendon Press.
- Freedman, Lawrence. 1981. *The evolution of nuclear strategy*. London: Macmillan.
- Frei, Daniel. 1983. *The risks of unintentional nuclear war*. London: Croom Helm for the U.N. Institute for Disarmament Research.
- French, Peter A. 1979. The corporation as a moral person. *American Philosophical Quarterly* 16: 207-15.
- French, Peter A. 1984. *Collective and corporate responsibility*. New York: Columbia University Press.
- Fried, Charles. 1978. *Right and wrong*. Cambridge, Mass.: Harvard University Press.
- Fried, Charles. 1981. *Contract as promise*. Cambridge, Mass.: Harvard University Press.
- Friedman, Milton. 1947. Lerner on the economics of control. *Journal of Political Economy* 55: 310-1.
- Friedman, Milton, and Rose Friedman. 1980. *Free to choose*. Harmondsworth: Penguin.
- Fuller, Lon. 1964. *The morality of law*. New Haven, Conn.: Yale University Press.
- Galanter, Marc. 1974. Why the "haves" come out ahead: speculations on the limits of legal change. *Law and Society Review* 9: 95-160.
- Garfinkel, Irwin, and Marieka M. Klawitter. 1990. The effect of routine withholding of child support collections. *Journal of Policy Analysis and Management* 9: 155-77.
- Gauthier, David. 1984. Deterrence, maximization and rationality. *Ethics* 94: 474-95. Reprinted in Hardin et al. 1985, pp. 99-120.
- Georgescu-Roegen, Nicholas. 1954. Choice, expectations and measurability. *Quarterly Journal of Economics* 69: 503-34.
- Gibbard, Allan. 1986. Interpersonal comparisons: preference, good and the intrinsic reward of a life. In Elster and Hylland 1986, pp. 165-94.
- Godwin, William. 1971. *Enquiry concerning political justice*. Oxford: Clarendon Press; originally published 1793.
- Goldfarb, Robert S. 1980. Compensating victims of policy change. *Regulation* (September-October): 22-30.

References

- Goodin, Robert E. 1975. Cross-cutting cleavages and social conflict. *British Journal of Political Science* 5: 316-19.
- Goodin, Robert E. 1976. *The politics of rational man*. London: Wiley.
- Goodin, Robert E. 1980. *Manipulatory politics*. New Haven, Conn.: Yale University Press.
- Goodin, Robert E. 1982a. Discounting discounting. *Journal of Public Policy* 2: 53-72.
- Goodin, Robert E. 1982b. *Political theory and public policy*. Chicago: University of Chicago Press.
- Goodin, Robert E. 1983. The ethics of destroying irreplaceable assets. *International Journal of Environmental Studies* 21: 55-66.
- Goodin, Robert E. 1985a. Disarming nuclear apologists. *Inquiry* 28: 153-76.
- Goodin, Robert E. 1985b. Erring on the side of kindness in social welfare policy. *Policy Sciences* 18: 141-56.
- Goodin, Robert E. 1985c. *Protecting the vulnerable*. Chicago: University of Chicago Press.
- Goodin, Robert E. 1988. *Reasons for welfare*. Princeton, N.J.: Princeton University Press.
- Goodin, Robert E. 1989a. The ethics of smoking. *Ethics* 99: 575-624.
- Goodin, Robert E. 1989b. *No smoking: the ethical issues*. Chicago: University of Chicago Press.
- Goodin, Robert E. 1989c. Stars to steer by: the political impact of moral values. *Journal of Public Policy* 9: 241-59.
- Goodin, Robert E. 1991a. Actual preferences, actual people. *Utilitas* 3: 113-19.
- Goodin, Robert E. 1991b. Permissible paternalism: in defense of the nanny state. *The Responsive Community* 1: 42-51.
- Goodin, Robert E. 1991c. Utility and the good. In Singer 1991, pp. 241-8.
- Goodin, Robert E. 1992a. *Green political theory*. Oxford: Polity.
- Goodin, Robert E. 1992b. *Motivating political morality*. Oxford: Blackwell.
- Goodin, Robert E., and Julian Le Grand, with John Dryzek, Diane Gibson, Russell Hanson, Robert H. Haveman, and David Winter. 1987. *Not only the poor: the middle classes and the welfare state*. London: Allen & Unwin.
- Goodin, Robert E., and Philip Pettit, eds. 1993. *A companion to contemporary political philosophy*. Oxford: Blackwell.
- Goodin, Robert E., and Andrew Reeve, eds. 1989. *Liberal neutrality*. London: Routledge.
- Goodin, Robert E., and K. W. S. Roberts. 1975. The ethical voter. *American Political Science Review* 69: 926-8.
- Grant, J. A. C. 1931. The "higher law" background of the law of eminent domain. *Wisconsin Law Review* 6: 67-85.
- Gray, John. 1981. John Stuart Mill on liberty, utility and rights. Pp. 80-116 in J. R. Pennock and J. W. Chapman, eds., *Nomos XXIII: human rights*. New York: New York University Press.
- Gregory, Robin, and Tim McDaniels. 1987. Valuing environmental losses: what promise does the right measure hold? *Policy Sciences* 20: 11-26.
- Griffin, James. 1977. Are there incommensurable values? *Philosophy and Public Affairs* 7: 38-59.
- Gross, B. M. 1953. *The legislative struggle*. New York: McGraw-Hill.

References

- Grotius, Hugo. 1925. *On the law of war and peace*. Trans. F. W. Kelsey. Oxford: Clarendon Press; originally published 1625.
- Halberstam, David. 1969. *The best and the brightest*. New York: Random House.
- Halévy, Elie. 1928. *The growth of philosophic radicalism*. Trans. Mary Morris. London: Faber & Faber.
- Hamblin, Charles. 1959. The modal "probably." *Mind* 68: 234-40.
- Hamermesh, Daniel S. 1979. Entitlement effects, unemployment insurance and employment decisions. *Economic Inquiry* 17: 317-32.
- Hampshire, Stuart. 1978. Morality and pessimism. Leslie Stephen Lecture, University of Cambridge, 1972. Reprinted pp. 1-22 in Stuart Hampshire, ed., *Public and private morality*. Cambridge: Cambridge University Press.
- Handler, Joel F. 1979. *Protecting the social service client*. New York: Academic Press.
- Hanson, Russell L. 1987. The expansion and contraction of the American welfare state. In Goodin and Le Grand 1987, pp. 169-202.
- Hardin, Russell. 1982. *Collective action*. Baltimore: Johns Hopkins University Press.
- Hardin, Russell. 1983. Unilateral versus mutual disarmament. *Philosophy and Public Affairs* 12: 236-54.
- Hardin, Russell. 1986. The utilitarian logic of liberalism. *Ethics* 97: 47-74.
- Hardin, Russell. 1988. *Morality within the limits of reason*. Chicago: University of Chicago Press.
- Hardin, Russell; John J. Mearsheimer; Gerald Dworkin; and Robert E. Goodin, eds. 1985 *Nuclear deterrence: ethics and strategy*. Chicago: University of Chicago Press.
- Hare, R. M. 1957. Reasons of state. Pp. 8-23 in R. M. Hare, *Applications of moral philosophy*. London: Macmillan, 1972.
- Hare, R. M. 1981. *Moral thinking*. Oxford: Clarendon Press.
- Harris, Donald A. 1974. The legal system of compensation for death and personal injury suffered in accidents. In Lees and Shaw 1974, pp. 30-55.
- Harris, Donald; Mavis Maclean; Hazel Genn; Sally Lloyd-Bostock; Paul Fenn; Peter Corfield; and Yvonne Brittan. 1994. *Compensation and support for illness and injury*. Oxford: Clarendon Press.
- Harris, John. 1975. The survival lottery. *Philosophy* 50: 81-7.
- Harrod, Roy F. 1936. Utilitarianism revisited. *Mind* 45: 137-56.
- 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? *American Political Science Review* 69: 594-606.
- Harsanyi, John C. 1977a. *Rational behavior and bargaining equilibrium in games and social situations*. Cambridge: Cambridge University Press.
- Harsanyi, John C. 1977b. Rule utilitarianism and decision theory. *Erkenntnis* 11: 25-53.
- Hart, H. L. A. 1955. Are there any natural rights? *Philosophical Review* 64: 175-91.
- Hart, H. L. A. 1961. *The concept of law*. Oxford: Clarendon Press.
- Hart, H. L. A. 1968. *Punishment and responsibility*. Oxford: Clarendon Press.
- Hart, H. L. A. 1983. Between utility and rights. Pp. 77-98 in Alan Ryan, ed.,

References

- The idea of freedom*. Oxford: Clarendon Press, 1979. Reprinted pp. 198-222 in H. L. A. Hart, *Essays in jurisprudence and philosophy*. Oxford: Clarendon Press.
- Hart, H. L. A., and Tony Honoré. 1985. *Causation in the law*. 2nd ed. Oxford: Clarendon Press; originally published 1959.
- Harvard Nuclear Study Group (Scott D. Sagan et al.). 1983. *Living with nuclear weapons*. Cambridge, Mass.: Harvard University Press.
- Harveman, Robert H., Victor Halberstadt, and Richard V. Burkhauser. 1984. *Public policy toward disabled workers*. Ithaca, N.Y.: Cornell University Press.
- Hayek, Friedrich A. 1945. The use of knowledge in society. *American Economic Review* 35: 519-30.
- Hayek, Friedrich A. 1960. *The constitution of liberty*. London: Routledge & Kegan Paul.
- Hayter, Theresa. 1981. *The creation of world poverty*. London: Pluto Press.
- Head, John G. 1974. *Public goods and public welfare*. Durham, N.C.: Duke University Press.
- Heil, John. 1988. Privileged access. *Mind* 42: 238-51.
- Held, Virginia. 1970. Can random collections of individuals be morally responsible? *Journal of Philosophy* 67: 471-81.
- Held, Virginia. 1972. Moral responsibility and collective action. Pp. 101-18 in Peter A. French, ed., *Individual and collective responsibility*. Cambridge, Mass.: Schenkman.
- Held, Virginia; Sidney Morgenbesser; and Thomas Nagel, eds. 1974. *Philosophy, morality and international affairs*. New York: Oxford University Press.
- Heyd, David. 1982. *Supererogation*. Cambridge: Cambridge University Press.
- Hicks, John R. 1939. The foundations of welfare economics. *Economic Journal*, 49: 696-712.
- Hill, Thomas E., Jr. 1979. Symbolic protest and calculated silence. *Philosophy and Public Affairs* 9: 83-102.
- Hirsch, Fred. 1976. *Social limits to growth*. Cambridge, Mass.: Harvard University Press.
- Hitch, Charles J., and Roland N. McKean. 1960. *The economics of defense in the nuclear age*. Cambridge, Mass.: Harvard University Press.
- Hochman, Harold M. 1974. Rule change and transitional equity. Pp. 320-41 in H. M. Hochman and George E. Peterson, eds., *Redistribution through public choice*. New York: Columbia University Press.
- Hodgson, D. H. 1967. *The consequences of utilitarianism*. Oxford: Clarendon Press.
- Hollis, Martin. 1977. *Models of man*. Cambridge: Cambridge University Press.
- Hollis, Martin. 1981. Economic man and original sin. *Political Studies* 29: 167-80.
- Holmes, Oliver Wendell. 1881. *The common law*. Boston: Little, Brown.
- Holmes, Oliver Wendell. 1897. The path of the law. *Harvard Law Review* 10: 457-78.
- Honoré, Tony. 1982. The human community and the principle of majority rule. Pp. 147-60 in Eugene Kamenka, ed., *Community as a social ideal*. London: Edward Arnold.

References

- Honoré, Tony. 1993. The dependence of morality on law. *Oxford Journal of Legal Studies* 13: 1-17.
- Hoole, Francis W., and Dina A. Zinnes, eds. 1976. *Quantitative international politics: an analysis*. New York: Praeger.
- Hume, David 1739. *A treatise of human nature*. London: John Noon.
- Hume, David. 1777. *An enquiry concerning the principles of morals*. London: Cadell.
- Hyman, H. H. 1942. The psychology of status. *Archives of Psychology*, no. 259.
- Isen, Terence G. 1980. *Accident compensation*. London: Croom Helm.
- Jackson, Frank. 1988. Group morality. Pp. 91-110 in P. Pettit, R. Sylvan and J. Norman, eds., *Metaphysics and morality*. Oxford: Blackwell.
- Jackson, Kenneth. 1985. *The crabgrass frontier*. New York: Oxford University Press.
- Jefferson, Thomas. 1801. Manual of parliamentary practice. Pp. 700-54 in G. F. Harrison and J. P. Coder, eds., *Senate manual*. 89th Congress, 1st Session. Washington, D.C.: Government Printing Office, 1965.
- Jervis, Robert. 1979. Deterrence theory revisited. *World Politics* 31: 289-324.
- Jevons, Stanley. 1911. *The theory of political economy*. 4th ed. London: Macmillan.
- Johnson, Norman. 1987. *The welfare state in transition*. Brighton: Wheatsheaf.
- Johnson, Norman. 1990. Problems for the mixed economy of welfare. In Ware and Goodin 1990, pp. 145-64.
- Jones, Peter. 1982. Freedom and the redistribution of resources. *Journal of Social Policy* 11: 217-38.
- Kagan, Shelly. 1987. Donagan and the sins of consequentialism. *Canadian Journal of Philosophy* 17: 643-54.
- Kagan, Shelly. 1989. *The limits of morality*. Oxford: Clarendon Press.
- Kahneman, Daniel; Paul Slovic; and Amos Tversky, eds. 1982. *Judgment under uncertainty: heuristics and biases*. Cambridge: Cambridge University Press.
- Kahneman, Daniel, and Amos Tversky. 1979. Prospect theory: an analysis of decision under risk. *Econometrica* 47: 163-91.
- Kaldor, Nicholas. 1939. Welfare propositions of economics and interpersonal comparisons of utility. *Economic Journal* 49: 549-52.
- Kaldor, Nicholas. 1946-7. Community indifference: a comment. *Review of Economic Studies* 14: 49.
- Kaplow, Louis. 1986. An economic analysis of legal transitions. *Harvard Law Review* 99: 509-618.
- Kavka, Gregory S. 1980. Deterrence, utility and rational choice. *Theory and Decision* 12: 41-60.
- Keeny, Spurgeon M., Jr., and Wolfgang K. H. Panofsky. 1981-2. MAD versus NUTS. *Foreign Affairs* 60: 287-304.
- Kennan, George F. 1970. To prevent a world wasteland. *Foreign Affairs* 48: 401-13.
- Kiewiet, D. Roderick. 1983. *Micropolitics and macroeconomics*. Chicago: University of Chicago Press.
- Kinder, Donald R., and D. Roderick Kiewiet. 1981. Sociotropic politics: the American case. *British Journal of Political Science* 11: 129-61.

References

- Kleinig, John. 1976. Good samaritanism. *Philosophy and Public Affairs* 5: 382-407.
- Kleinig, John. 1982. The ethics of consent. *Canadian Journal of Philosophy (Supplement)* 8: 91-118.
- Knetsch, Jack L., and J. A. Sinden. 1984. Willingness to pay and compensation demanded: experimental evidence of an unexpected disparity in measures of value. *Quarterly Journal of Economics* 99: 507-21.
- Kronman, Anthony. 1980. Contract law and distributive justice. *Yale Law Journal* 89: 472-511.
- Kuperberg, Mark, and Charles Beitz, eds. 1983. *Law, economics and philosophy*. Totowa, N.J.: Rowman & Allanheld.
- Kymlicka, Will. 1989. *Liberalism, community and culture*. Oxford: Clarendon Press.
- Lackey, Douglas P. 1982. Morals and missiles: a utilitarian look at nuclear deterrence. *Philosophy and Public Affairs* 11: 189-231.
- Lackey, Douglas P. 1984. *Moral principles and nuclear weapons*. Totowa, N.J.: Rowman & Allanheld.
- Lakatos, Irme, and Alan Musgrave, eds. 1970. *Criticism and the growth of knowledge*. Cambridge: Cambridge University Press.
- Lancaster, Kelvin J. 1966. A new approach to consumer theory. *Journal of Political Economy* 74: 132-57.
- Lancaster, Kelvin J. 1971. *Consumer demand: a new approach*. New York: Columbia University Press.
- Land, Hillary. 1978. Who cares for the family? *Journal of Social Policy* 7: 257-84.
- Lave, C. A., and James G. March. 1975. *An introduction to models in the social sciences*. New York: Harper & Row.
- Layard, Richard, ed. 1974. *Cost-benefit analysis*. Rev. ed. Harmondsworth: Penguin.
- Le Grand, Julian. 1982. *The strategy of equality*. London: Allen & Unwin.
- Le Grand, Julian. 1990. Equity versus equality: the elusive trade-off. *Ethics* 100: 554-68.
- Lees, Dennis, and Stella Shaw, eds. 1974. *Impairment, disability and handicap*. London: Heinemann for the SSRC.
- Levine, Henry D. 1977. Some things to all men: the politics of Cruise missile development. *Public Policy* 25: 118-68.
- Lewis, David. 1969. *Convention*. Oxford: Blackwell.
- Lipsky, Michael. 1984. Bureaucratic disenitment in social welfare programs. *Social Service Review* 58: 3-27.
- Little, I. M. D. 1952. Social choice and individual values. *Journal of Political Economy* 60: 422-32.
- Little, I. M. D. 1957. *A critique of welfare economics*. 2nd ed. Oxford: Clarendon Press.
- Lively, Jack, and John Rees, eds. 1978. *Utilitarian logic and politics*. Oxford: Clarendon Press.
- Llewellyn, Karl. 1960. *The bramble bush*. 2nd ed. Dobbs Ferry, N.Y.: Oceana.
- Locke, John. 1946. *Letter concerning toleration*. Pp. 125-67 in John Locke, *Second*

References

- treatise of government*, ed. J.W. Gough. Oxford: Blackwell; originally published 1689.
- Luce, R. Duncan, and Howard Raiffa. 1957. *Games and decisions*. New York: Wiley.
- Lukes, Steven. 1993. Five fables about human rights: what would it be like if ... *Dissent* (Fall): 427-37.
- Lyons, David. 1965. *The forms and limits of utilitarianism*. Oxford: Clarendon Press.
- Maass, Arthur. 1966. Benefit-cost analysis: its relevance to public investment decisions. *Quarterly Journal of Economics* 80: 208-26.
- Mabbott, J. D. 1939. Punishment. *Mind* 48: 152-67.
- Mabbott, J. D. 1956. Interpretations of Mill's "Utilitarianism." *Philosophical Quarterly* 6: 115-29.
- Macaulay, T. B. 1978. Mill's essay on government: utilitarian logic and politics. *Edinburgh Review* 97 (March 1829): article 7. Reprinted in Lively and Rees 1978, pp. 99-129.
- MacCallum, Gerald C. 1966. Legislative intent. *Yale Law Journal* 75: 754-87.
- MacCormack, Geoffrey. 1973. Revenge and compensation in early law. *American Journal of Comparative Law* 21: 69-85.
- MacCormick, Neil. 1982. *Legal right and social democracy*. Oxford: Clarendon Press.
- MacIntyre, Alisdair. 1967. Egoism and altruism. In Edwards 1967, vol. 2, pp. 462-66.
- MacIntyre, Alisdair. 1984. *Is patriotism a virtue?* Lindley lecture, University of Kansas, March 26 1984. Lawrence, Kansas: University of Kansas.
- MacIver, Robert M. 1940. The imputation of motives. *American Journal of Sociology* 46: 1-12.
- Mackie, J. L. 1955. Responsibility and language. *Australasian Journal of Philosophy* 33: 143-59.
- MacRae, Duncan, Jr. 1976. *The social function of social science*. New Haven, Conn.: Yale University Press.
- Mansbridge, Jane J., ed. 1990. *Beyond self-interest*. Chicago: University of Chicago Press.
- Mansbridge, Jane J., and Susan Moller Okin. 1993. Feminism. In Goodin and Pettit 1993, pp. 269-90.
- March, James G., and Johan P. Olsen. 1976. *Ambiguity and choice in organizations*. Bergen: Universitetsforlaget.
- Marglin, Stephen A. 1963. The social rate of discount and the optimal rate of investment. *Quarterly Journal of Economics* 77: 95-111.
- Margolis, Howard. 1982. *Selfishness, altruism and rationality*. Cambridge: Cambridge University Press.
- Marshall, Alfred. 1920. *Principles of economics*. 8th ed. London: Macmillan.
- Marx, Karl, and Friedrich Engels. 1958. *Selected works*. Moscow: Foreign Languages Publishing House.
- Mathews, Jessica Tuchman. 1989. Redefining security. *Foreign Affairs* 68 (2): 162-77.
- May, Ernest. 1971. *The law, privileges, proceedings and usage of Parliament*, ed. B. Cocks. 18th ed. London: Butterworth.

References

- Mayo, H. B. 1960. *An introduction to democratic theory*. New York: Oxford University Press.
- McCloskey, Herbert A. 1964. Consensus and ideology in American politics. *American Political Science Review* 58: 361-82.
- McCloskey, Herbert A., and Alida Brill. 1983. *Dimensions of tolerance*. New York: Russell Sage Foundation.
- McGraw, Kathleen M. 1990. Avoiding blame: an experimental investigation of political excuses and justifications. *British Journal of Political Science* 20: 119-42.
- McKinsey, Michael. 1981. Obligations to the starving. *Notis* 15: 309-23.
- Meade, James E. 1972. The theory of labour-managed firms and of profit sharing. *Economic Journal* 82: 402-28.
- Mellema, Gregory. 1984. On being fully responsible. *American Philosophical Quarterly* 21: 189-93.
- Michelman, Frank I. 1967. Property, utility, and fairness: comments on the ethical foundations of "just compensation" law. *Harvard Law Review* 80: 1165-1258.
- Mill, James. 1992. *Essays on government, jurisprudence, [etc.]*. Reprinted as *James Mill: political writings*, ed. Terence Ball. Cambridge: Cambridge University Press; originally published 1823.
- Mill, John Stuart. 1848. *Principles of political economy*. London: Parker & Son.
- Mill, John Stuart. 1962a. Essay on Bentham. *London and Westminster Review* (August 1838). Reprinted in Warnock 1962, pp. 78-125.
- Mill, John Stuart. 1962b. *Utilitarianism*. London: Parker & Son, 1863. Reprinted in Warnock 1962, pp. 251-321.
- Mill, John Stuart. 1975a. *Considerations on representative government*. In Wollheim 1975, pp. 142-423; originally published 1861.
- Mill, John Stuart. 1975b. *On liberty*. In Wollheim 1975, pp. 5-141; originally published 1859.
- Mill, John Stuart. 1975c. *The subjection of women*. In Wollheim 1975, pp. 425-58; originally published 1869.
- Miller, David. 1981. Market neutrality and the failure of cooperatives. *British Journal of Political Science* 11: 309-21.
- Miller, David. 1988. The ethical significance of nationality. *Ethics* 98: 647-62.
- Mills, C. Wright. 1940. Situated actions and vocabularies of motive. *American Sociological Review* 5: 904-13.
- Mises, Ludwig von. 1949. *Human action*. London: William Hodge.
- Mishan, E. J. 1967. Pareto optimality and the law. *Oxford Economic Papers* 19: 255-87.
- Mishan, E. J. 1971. The post-war literature on externalities. *Journal of Economic Literature* 9: 1-28.
- Mishan, E. J. 1974. What is wrong with Roskill? In Layard 1974, pp. 452-72; originally published 1970.
- Mondale, Walter F., Lloyd Bentsen, and Abraham Ribicoff. 1974. Senate report number 93-1356: Social Security Amendments of 1974. *United States Code: Congressional and Administrative News* (93rd Congress, 2nd Session) 4: 8133-60.

References

- Montreal Protocol. 1987. Protocol on substances that deplete the ozone layer. *International Legal Materials* 26: 1550-61.
- Moon, J. Donald, ed. 1988. *Responsibility, rights and welfare*. Westport, Conn.: Greenwood Press.
- Moore, G. E. 1903. *Principia ethica*. Cambridge: Cambridge University Press.
- Morris, Herbert. 1981. A paternalistic theory of punishment. *American Philosophical Quarterly* 18: 263-71.
- Murray, Charles A. 1984. *Losing ground: American social policy, 1950-1980*. New York: Basic Books.
- Musgrave, Richard A. 1968. Public expenditures. In Sills 1968, vol. 13, pp. 156-64.
- Myrdal, Gunnar. 1944. *An American dilemma*. London: Harper.
- Nagel, Thomas. 1978. Ruthlessness in public life. Pp. 75-92 in Stuart Hampshire, ed., *Public and private morality*. Cambridge: Cambridge University Press.
- Nagel, Thomas. 1986. *The view from nowhere*. New York: Oxford University Press.
- Nisbett, Richard E., and Timothy DeCamp Wilson. 1977. Telling more than we can know: verbal reports on mental processes. *Psychological Review* 84: 231-59.
- Nonet, Philippe. 1969. *Administrative justice*. New York: Russell Sage Foundation.
- Norman, Richard. 1971. *Reasons for action*. Oxford: Blackwell.
- Novack, Michael, chairman. 1987. *The new consensus on family and welfare*. Report of the Working Seminar on Family and American Welfare Policy. Washington, D.C.: American Enterprise Institute.
- Nozick, Robert. 1968. Moral complications and moral structures. *Natural Law Forum* 13: 1-50.
- Nozick, Robert. 1972. Coercion. Pp. 110-35 in P. Laslett, W. G. Runciman, and Q. Skinner, eds., *Philosophy, politics and society*. 4th series. Oxford: Blackwell.
- Nozick, Robert. 1974. *Anarchy, state and utopia*. Oxford: Blackwell.
- Nozick, Robert. 1981. *Philosophical explanations*. Cambridge, Mass.: Harvard University Press.
- Offe, Claus. 1992. A non-productivist design for social policies. In Van Parijs 1992, pp. 61-80.
- O'Hare, Michael. 1977. "Not on my block you don't": facility siting and the strategic importance of compensation. *Public Policy* 25: 407-58.
- Okun, Arthur. 1975. *Equality and efficiency: the big tradeoff*. Washington, D.C.: Brookings Institution.
- Olsen, Johan P. 1972. Public policy-making and theories of organizational choice. *Scandinavian Political Studies* 7: 45-62.
- O'Neill, Onora. 1985. Between consenting adults. *Philosophy and Public Affairs* 14: 252-77.
- Oppenheim, L. 1955. *International law: a treatise*, ed. H. Lauterpact. London: Longman.
- Organisation for Economic Co-operation and Development (OECD). 1981. *The welfare state in crisis*. Paris: OECD.

References

- Page, Talbot. 1978. A generic view of toxic chemicals and similar risks. *Ecology Law Quarterly* 7: 207-44.
- Pahl, Jan. 1983. The allocation of money and the structuring of inequality within marriage. *Sociological Review* 31: 237-62.
- Palme, Joakim. 1990. Models of old-age pensions. In Ware and Goodin 1990, pp. 104-25.
- Pareto, Vilfredo 1971. *Manual of political economy*. Trans. Ann S. Schwier. New York: Augustus M. Kelly; originally published 1927.
- Parfit, Derek. 1973. Later selves and moral principles. Pp. 137-69 in Alan Montefiore, ed., *Philosophy and personal relations*. London: Routledge & Kegan Paul.
- Parfit, Derek. 1976. Lewis, Perry and what matters. Pp. 91-108 in A. O. Rorty, ed., *The identities of persons*. Berkeley: University of California Press.
- Parfit, Derek. 1984. *Reasons and persons*. Oxford: Clarendon Press.
- Partington, Martin. 1978. *Claim in time*. London: Frances Pinter.
- Paskins, Barrie, and Michael Dockrill. 1979. *The ethics of war*. London: Duckworth.
- Payne, Keith B., and Colin S. Gray. 1984. Nuclear policy and the defensive transition. *Foreign Affairs* 62: 820-42.
- Pears, David. 1984. *Motivated irrationality*. Oxford: Clarendon Press.
- Pechman, Joseph A., Henry J. Aaron, and Michael K. Taussig. 1968. *Social security: perspectives for reform*. Washington, D.C.: Brookings Institution.
- Peters, R. S. 1956. Motives and motivation. *Philosophy* 31: 117-30.
- Pettit, Philip. 1986. Free riding and foul dealing. *Journal of Philosophy* 83: 361-79.
- Pettit, Philip. 1991. Consequentialism. In Singer 1991, pp. 230-40.
- Pettit, Philip, and Robert E. Goodin. 1986. The possibility of special duties. *Canadian Journal of Philosophy* 16: 651-76.
- Piachaud, David. 1981. Peter Townsend and the Holy Grail. *New Society* 54 (982): 419-21.
- Piachaud, David. 1982. Patterns of income and expenditure within families. *Journal of Social Policy* 11: 469-82.
- Pigou, A. C. 1932. *The economics of welfare*. 4th ed. London: Macmillan.
- Pigou, A. C. 1951. Some aspects of economic welfare. *American Economic Review* 41: 287-302.
- Plamenatz, John. 1958. *The English utilitarians*. Oxford: Blackwell.
- Plamenatz, John. 1973. *Democracy and illusion*. London: Longman.
- Plant, Raymond. 1988. Needs, agency, and welfare rights. In Moon 1988, pp. 55-75.
- Pringle, Peter, and William Arkin. 1983. *SIOP*. London: Sphere Books.
- Prior, A. N. 1956. Symposium on the consequences of actions: I. *Proceedings of the Aristotelian Society (Supplement)* 30: 91-9.
- Prosser, William. 1964. *Handbook of the law of torts*. 3rd ed. St. Paul, Minn.: West.
- Prosser, William. 1965. *Restatement (second) of the law of torts*. 3rd ed. St. Paul, Minn.: American Law Institute.
- Prosser, William. 1971. *Handbook of the law of torts*. 4th ed. St. Paul, Minn.: West.

References

- Prosser, William, and J. W. Wade. 1979. *Restatement (second) of the law of torts*. 4th ed. St. Paul, Minn.: West for the American Law Institute.
- Prothro, James W., and Charles M. Grigg. 1960. Fundamental principles of democracy: bases of agreement and disagreement. *Journal of Politics* 22: 276-94.
- Pufendorf, Samuel. 1717. *Of the law of nature and nations*. Trans. B. Kennett. 3rd ed. London: Sare.
- Raiffa, Howard. 1968. *Decision analysis*. Reading, Mass.: Addison-Wesley.
- Raphael, D. D. 1956. Symposium on the consequences of actions: II. *Proceedings of the Aristotelian Society (Supplement)* 30: 100-19.
- Rawls, John. 1955. Two concepts of rules. *Philosophical Review* 64: 3-32.
- Rawls, John. 1971. *A theory of justice*. Cambridge, Mass.: Harvard University Press.
- Raz, Joseph. 1975. *Practical reasons and norms*. London: Hutchinson.
- Raz, Joseph. 1986. *The morality of freedom*. Oxford: Clarendon Press.
- Reeve, Andrew. 1990. Individual choice and the retreat from utilitarianism. Pp. 98-119 in Lincoln Allison, ed., *The utilitarian response*. London: Sage.
- Regan, Donald H. 1980. *Utilitarianism and co-operation*. Oxford: Clarendon Press.
- Reich, Charles A. 1963. Midnight welfare searches and the Social Security Act. *Yale Law Journal* 72: 1347-60.
- Riley, Patrick. 1986. *The general will before Rousseau*. Princeton, N.J.: Princeton University Press.
- Robbins, Lionell. 1932. *Essay on the nature and significance of economic science*. London: Macmillan.
- Robbins, Lionell. 1938. Interpersonal utility comparisons. *Economic Journal* 48: 635-41.
- Rorty, A. O., ed. 1976. *The identities of persons*. Berkeley: University of California Press.
- Rose, Richard, and B. Guy Peters. 1979. *Can government go bankrupt?* London: Macmillan.
- Rosenberg, Gerald M. 1977. The protection of aliens from discriminatory treatment by the national government. *Supreme Court Review* 1977: 275-339.
- Ross, W. D. 1930. *The right and the good*. Oxford: Clarendon Press.
- Rothenberg, Jerome. 1968. Consumer sovereignty. In Sills 1968, vol. 3, pp. 326-35.
- Rousseau, Jean-Jacques. 1911. *Emile*. Trans. Barbara Foxley. London: Dent; originally published 1762.
- Rousseau, Jean-Jacques. 1973. The social contract. Pp. 164-278 in *The social contract and discourses*, trans. G. D. H. Cole. New ed. London: Everyman/Dent; originally published 1762.
- Runciman, W. G. 1966. *Relative deprivation and social justice*. London: Routledge & Kegan Paul.
- Runciman, W. G., and A. K. Sen. 1965. Games, justice and the general will. *Mind* 74: 554-62.
- Ryle, Gilbert. 1949. *The concept of mind*. London: Hutchinson.
- Sagan, Carl. 1983-4. Nuclear war and climatic consequence: some policy implications. *Foreign Affairs* 62: 257-92.

References

- Sagoff, Mark. 1974. On preserving the natural environment. *Yale Law Journal* 84: 205-67.
- Sagoff, Mark. 1978. On restoring and reproducing art. *Journal of Philosophy* 75: 453-70.
- Sagoff, Mark. 1988. *The economy of the earth*. Cambridge: Cambridge University Press.
- St. Leonards (Lord Chancellor). 1852. *Dundee Harbour Trustees v. Dougall*. Vol. 1, pp. 14-23, in James Paterson, ed., *Reports of Scotch appeals in the House of Lords, 1851-61*. Edinburgh: William Green & Sons, 1985.
- Sandel, Michael J. 1982. *Liberalism and the limits of justice*. Cambridge: Cambridge University Press.
- Sandel, Michael J., ed. 1984. *Liberalism and its critics*. Oxford: Blackwell.
- Sax, Joseph L. 1971. Takings, private property and public rights. *Yale Law Journal* 81: 149-86.
- Saylor, Taylor Reese. 1977. A new legislative mandate for American food aid. Pp. 199-211 in Peter G. Brown and Henry Shue, eds., *Food policy*. New York: Free Press.
- Scanlon, T. M. 1975. Preference and urgency. *Journal of Philosophy* 72: 655-9.
- Scheffler, Samuel. 1982. *The rejection of consequentialism*. Oxford: Clarendon Press.
- Schlesinger, Arthur M., Jr. 1981. Unlearning the lessons of Pearl Harbor. *Washington Post* (December 6): C1, C3.
- Schoeck, Helmut. 1969. *Envy*. Trans. Michael Glenny and Betty Ross. New York: Harcourt, Brace & World; originally published 1966.
- Schuck, Peter H. 1984. The transformation of immigration law. *Columbia Law Review* 84: 1-90.
- Scitovsky, Tibor. 1941. A note on welfare propositions in economics. *Review of Economic Studies* 9: 77-88.
- Sears, David O., R. R. Lau, Tom R. Tyler, and H. M. Allen, Jr. 1980. Self-interest vs. symbolic politics in policy attitudes and presidential voting. *American Political Science Review* 74: 670-84.
- Self, Peter. 1975. *Econocrats and the policy process*. London: Macmillan.
- Sen, Amartya. 1957. A note on Tinbergen on the optimum rate of saving. *Economic Journal* 67: 745-48.
- Sen, Amartya. 1967. Isolation, assurance and the social rate of discount. *Quarterly Journal of Economics* 81: 112-24.
- Sen, Amartya. 1970. *Collective choice and social welfare*. San Francisco: Holden-Day.
- Sen, Amartya. 1973. Behaviour and the concept of preference. *Economica* 40: 241-59. Reprinted in Sen 1982a, pp. 54-73.
- Sen, Amartya. 1974a. Choice, ordering and morality. Pp. 54-67 in Stefan Körner, ed., *Practical reason*. Oxford: Blackwell.
- Sen, Amartya. 1974b. Rawls versus Bentham: an axiomatic examination of the pure distribution problem. *Theory and Decision* 4: 301-10.
- Sen, Amartya. 1976. Liberty, unanimity and rights. *Economica* 43: 217-45. Reprinted in Sen 1982a, pp. 291-326.
- Sen, Amartya. 1977a. Rational fools: a critique of the behavioral foundations

References

- of economic theory. *Philosophy and Public Affairs* 6: 317-44. Reprinted in Sen 1982a, pp. 84-106.
- Sen, Amartya. 1977b. Starvation and exchange entitlements: a general approach and its application to the Great Bengal Famine. *Cambridge Journal of Economics* 1: 33-59.
- Sen, Amartya. 1979a. Personal utilities and public judgments: or what's wrong with welfare economics? *Economic Journal* 89: 537-58. Reprinted in Sen 1982a, pp. 327-52.
- Sen, Amartya. 1979b. Utilitarianism and welfarism. *Journal of Philosophy* 76: 463-89.
- Sen, Amartya. 1980-1. Plural utilities. *Proceedings of the Aristotelian Society* 81: 193-215.
- Sen, Amartya. 1981. *Poverty and famines*. Oxford: Clarendon Press.
- Sen, Amartya. 1982a. *Choice, welfare and measurement*. Oxford: Blackwell.
- Sen, Amartya. 1982b. Rights and agency. *Philosophy and Public Affairs* 11: 3-39.
- Sen, Amartya. 1983. Poor, relatively speaking. *Oxford Economic Papers* 35: 153-69.
- Sen, Amartya. 1985a. *Commodities and capabilities*. Amsterdam: North-Holland.
- Sen, Amartya. 1985b. Well-being, agency and freedom. *Journal of Philosophy* 82: 169-221.
- Sen, Amartya, and Bernard Williams, eds. 1982. *Utilitarianism and beyond*. Cambridge: Cambridge University Press.
- Shacknové, Andrew E. 1985. Who is a refugee? *Ethics* 95: 274-84.
- Shklar, Judith N. 1979. Let us not be hypocritical. *Daedalus* 108: 1-25. Reprinted in Shklar 1984, pp. 45-86.
- Shklar, Judith N. 1984. *Ordinary vices*. Cambridge, Mass.: Harvard University Press.
- Shubik, Martin. 1971. Games of status. *Behavioral Science* 16: 117-29.
- Shue, Henry. 1980. *Basic rights*. Princeton, N.J.: Princeton University Press.
- Shue, Henry. 1988. Mediating duties. *Ethics* 98: 687-704.
- Sidgwick, Henry. 1891. *The elements of politics*. London: Macmillan.
- Sidgwick, Henry. 1907. *The methods of ethics*. 7th ed. London: Macmillan; originally published 1874.
- Sills, David L., ed. 1968. *International encyclopedia of the social sciences*. 18 vols. London: Collier-Macmillan.
- Singer, Peter. 1972. Famine, affluence and morality. *Philosophy and Public Affairs* 1: 229-43.
- Singer, Peter. 1979. *Practical ethics*. Cambridge: Cambridge University Press.
- Singer, Peter, ed. 1991. *A companion to ethics*. Oxford: Blackwell.
- Skinner, B. F. 1948. *Walden II*. New York: Macmillan.
- Smart, J. J. C. 1967. Utilitarianism. In Edwards 1967, vol. 8, pp. 206-212.
- Smart, J. J. C. 1973. An outline of a system of utilitarian ethics. Pp. 1-73 in J. J. C. Smart and Bernard Williams, *Utilitarianism: for and against*. Cambridge: Cambridge University Press.
- Smith, Adam. 1790. *The theory of moral sentiments*. 6th ed. London: A. Strahan & T. Cadell.
- Sniderman, Paul M.; Philip E. Tetlock; James M. Glaser; Donald Philip Green;

References

- and Michael Hout. 1989. Principled tolerance and the American mass public. *British Journal of Political Science* 19: 47-67.
- Snyder, Jack. 1978. Rationality at the brink. *World Politics* 30: 345-65.
- Sohn, Louis B., and R. R. Baxter. 1961. Responsibility of states for injuries to the economic interests of allens (Harvard Law School draft convention). *American Journal of International Law* 55: 545-84.
- Stapleton, Jane. 1986. *Disease and the compensation debate*. Oxford: Clarendon Press.
- Steiner, Hillel. 1992. Three just taxes. In Van Parijs 1992, pp. 81-92.
- Stephen, James Fitzjames. 1874. *Liberty, equality, fraternity*. 2nd ed. London: Smith and Elder.
- Steuart, James. 1966. *An inquiry into the principles of political economy*, ed. A. S. Skinner. Edinburgh: Oliver & Boyd; originally published 1767.
- Stone, Deborah. 1984. *The disabled state*. Philadelphia: Temple University Press.
- Stone, Peter B. 1973. *Did we save the earth at Stockholm?* London: Earth Island.
- Streton, Paul, with Shahid Javed Burki, Mahbub ul Haq, Norman Hicks, and Frances Stewart. 1981. *First things first: meeting basic human needs in the developing countries*. New York: Oxford University Press for the World Bank.
- Suárez, Francisco. 1612. *Treatise on laws and God the lawgiver*. In Francisco Suárez, *Selections from three works*, ed. G. L. Williams, A. Brown, J. Waldron, and H. Davis. Oxford: Clarendon Press, 1944.
- Sunstein, Cass R. 1990. *After the rights revolution*. Cambridge, Mass.: Harvard University Press.
- Sunstein, Cass R. 1991. Preferences and politics. *Philosophy and Public Affairs* 20: 3-34.
- Taylor, C. C. W. 1969. Understanding a want. Pp. 244-61 in J. J. Macintosh and S. Coval, eds., *The business of reason*. London: Routledge & Kegan Paul.
- Taylor, Charles. 1976. Responsibility for self. In Rorty 1976, pp. 281-300.
- Taylor, E. 1971. *The House of Commons at work*. 8th ed. Harmondsworth: Penguin.
- Taylor, Michael, and Hugh Ward. 1982. Chickens, whales and lumpy public goods: alternative models of public-goods provision. *Political Studies* 30: 350-70.
- tenBroek, Jacobus. 1954. The impact of welfare law upon family law. *California Law Review* 42: 458-85.
- tenBroek, Jacobus. 1966. The right to live in the world: the disabled in the law of torts. *California Law Review* 54: 841-919.
- Thompson, Dennis F. 1987. *Political ethics and public office*. Cambridge, Mass.: Harvard University Press.
- Thomson, Judith Jarvis. 1977. *Acts and other events*. Ithaca, N.Y.: Cornell University Press.
- Thomson, Judith Jarvis. 1980. Rights and compensation. *Noûs* 14: 3-15.
- Thomson, Judith Jarvis. 1986. *Rights, restitution and risk*. Cambridge, Mass.: Harvard University Press.
- Titmuss, R. M. 1968. *Commitment to welfare*. London: Allen & Unwin.
- Titmuss, R. M. 1987a. The social division of welfare: some reflections on the

References

- search for equity. Reprinted in Abel-Smith and Titmuss 1987, pp. 39-59; originally published 1955.
- Titmuss, R. M. 1987b. Universal and selective social services. Reprinted in Abel-Smith and Titmuss 1987, pp. 128-40; originally published 1967.
- Tobin, James. 1970. On limiting the domain of inequality. *Journal of Law and Economics* 13: 363-78.
- Townsend, Peter. 1954. Measuring poverty. *British Journal of Sociology* 5: 130-7.
- Townsend, Peter. 1962. The meaning of poverty. *British Journal of Sociology* 13: 210-27.
- Townsend, Peter. 1979. *Poverty in the United Kingdom*. Harmondsworth: Penguin.
- Trachtenberg, Marc. 1985. Strategists, philosophers and the nuclear question. In Hardin et al. 1985, pp. 355-66.
- Tribe, Laurence H. 1972. Policy science: analysis or ideology. *Philosophy and Public Affairs* 2: 66-100.
- Tribe, Laurence H. 1978. *American constitutional law*. Mineola, N.Y.: Foundation Press.
- Tufte, Edward. 1978. *The political control of the economy*. Princeton, N.J.: Princeton University Press.
- Tullock, Gordon. 1978. Achieving deregulation: a public choice perspective. *Regulation* (November-December): 50-54.
- Tussman, Joseph, and Jacobus tenBroek. 1949. The equal protection of the laws. *California Law Review* 37: 341-81.
- Tversky, Amos. 1975. A critique of expected utility theory. *Erkenntnis* 9: 163-73.
- Tversky, Amos. 1981. Choice, preference and welfare: some psychological observations. Paper presented to colloquium on "Foundations of Social Choice Theory," Ustaoset, Norway.
- Tversky, Amos, and Daniel Kahneman. 1981. The framing of decisions and the psychology of choice. *Science* 211: 453-8.
- United Nations Conference on the Human Environment. 1972. Declaration on the Human Environment. *International Legal Materials* 11: 1416-21.
- United States Congress. Senate. 1965. Standing rules for the conduct of business in the Senate of the United States. Pp. 1-41 in G. F. Harrison and J. P. Coder, eds., *Senate manual*. 89th Congress, 1st Session. Washington, D.C.: Government Printing Office.
- United States Department of Health, Education, and Welfare (U.S. DHEW). 1978. *Social security programs throughout the world, 1977*. Social Security Administration, Office of Research and Statistics, Research Report no. 50. Washington, D.C.: Government Printing Office.
- Urmson, J. O. 1953. The interpretation of the moral philosophy of J. S. Mill. *Philosophical Quarterly* 3: 33-9.
- Urmson, J. O. 1958. Saints and heroes. Pp. 198-216 in A. I. Melden, ed., *Essays in moral philosophy*. Seattle: University of Washington Press.
- van der Veen, Robert, and Philippe Van Parijs. 1987. A capitalist road to communism. *Theory and Society* 15: 635-55.
- Van Parijs, Philippe. 1986. Basic income: a terminological note. Pp. 3-10 in

References

- Proceedings of the first international conference on basic income*. Louvain-la-Neuve, Belgium: Basic Income European Network.
- Van Parijs, Philippe. 1990. The second marriage of justice and efficiency. *Journal of Social Policy* 19: 1-25. Reprinted in Van Parijs 1992, pp. 215-43.
- Van Parijs, Philippe. 1991. Why surfers should be fed: the liberal case for an unconditional basic income. *Philosophy and Public Affairs* 20: 101-31.
- Van Parijs, Philippe, ed. 1992. *Arguing for basic income*. London: Verso.
- Vattel, Emerich de. 1863. *The law of nations, or the principles of natural law*. Trans. Joseph Chitty. Philadelphia: T. and J. W. Johnson; originally published 1758.
- Vienna Convention. 1985. Convention for the protection of the ozone layer. *International Legal Materials* 26 (1987): 1529-40.
- Waldron, Jeremy. 1986. Welfare and the images of charity. *Philosophical Quarterly* 36: 463-82. Reprinted in Waldron 1993a, pp. 225-49.
- Waldron, Jeremy. 1991. Homelessness and the issue of freedom. *UCLA Law Review* 39: 295-324. Reprinted in Waldron 1993a, pp. 309-38.
- Waldron, Jeremy. 1993a. *Liberal rights*. Cambridge: Cambridge University Press.
- Waldron, Jeremy. 1993b. Special ties and natural duties. *Philosophy and Public Affairs* 22: 3-30.
- Walsh, Clifford. 1987. Individual irrationality and public policy: in search of merit/demerit policies. *Journal of Public Policy* 7: 103-34.
- Walzer, Michael. 1973. Political action: the problem of dirty hands. *Philosophy and Public Affairs* 2: 160-80. Reprinted in Cohen et al. 1974, pp. 62-82.
- Walzer, Michael. 1977. *Just and unjust wars*. New York: Basic Books.
- Walzer, Michael. 1980. The moral standing of states. *Philosophy and Public Affairs* 9: 209-30. Reprinted in Beitz et al. 1985, pp. 217-37.
- Walzer, Michael. 1983. *Spheres of justice*. Oxford: Martin Robertson.
- Ward, Barbara, and Rene Dubos. 1972. *Only one earth*. Harmondsworth: Penguin.
- Ware, Alan, and Robert E. Goodin, eds. 1990. *Needs and welfare*. London: Sage.
- Warnock, Mary, ed. 1962. *John Stuart Mill: Utilitarianism; on liberty; essay on Bentham*. Cleveland: Meridian.
- Wasserstrom, Richard. 1975. Lawyers as professionals: some moral issues. *Human Rights* 5: 1-24.
- Wasserstrom, Richard. 1983. Roles and morality. Pp. 25-37 in David Luban, ed., *The good lawyer*. Totowa, N.J.: Rowman & Allanheld.
- Watson, Gary. 1975. Free agency. *Journal of Philosophy* 72: 205-20.
- Watson, Gary. 1977. Skepticism about weakness of will. *Philosophical Review* 86: 316-39.
- Watt, A. J. 1972. The intelligibility of wants. *Mind* 81: 553-61.
- Weaver, R. Kent. 1986. The politics of blame avoidance. *Journal of Public Policy* 6: 371-98.
- Weber, Max. 1919. Politics as a vocation. Pp. 77-128 in Hans Gerth and C. Wright Mills, eds., *From Max Weber: essays in sociology*. New York: Oxford University Press, 1946.
- Weinrib, E. J. 1980. The case for a duty to rescue. *Yale Law Journal* 90: 247-93.
- Weitzman, M. L. 1977. Is the price system or rationing more effective in get-

References

- ting a commodity to those who need it most? *Bell (now Rand) Journal of Economics* 8: 517-24.
- Wickstead, Philip H. 1933. *The common sense of political economy*. London: Routledge.
- Wiggins, David. 1985. Claims of need. Pp. 142-202 in Ted Honderich, ed., *Morality and objectivity*. London: Routledge & Kegan Paul.
- Wildavsky, Aaron. 1962. Nuclear clubs or nuclear wars? *Yale Review* 51: 345-62.
- Williams, Bernard. 1973a. A critique of utilitarianism. Pp. 75-150 in J. J. C. Smart and Bernard Williams, *Utilitarianism, for and against*. Cambridge: Cambridge University Press.
- Williams, Bernard. 1973b. *Problems of the self*. Cambridge: Cambridge University Press.
- Williams, Bernard. 1981. *Moral luck*. Cambridge: Cambridge University Press.
- Williams, Bernard. 1982. How to think skeptically about the bomb. *New Society*, vol. 62, no. 1044 (November 19): pp. 288-90.
- Williams, Bernard. 1983. Professional morality and its dispositions. Pp. 259-69 in David Luban, ed., *The good lawyer*. Totowa, N.J.: Rowman & Allanheld.
- Williams, Glanville. 1951. *Joint torts and contributory negligence*. London: Stevens.
- Williams, Glanville, and B. A. Hepple. 1984. *Foundations of the law of torts*. 2nd ed. London: Butterworth.
- Williston, S. 1932. *Restatement of the law of contracts*. St. Paul, Minn.: American Law Institute.
- Wirth, David A. 1989. Climate chaos. *Foreign Policy* 74: 3-22.
- Wohlstetter, Roberta. 1962. *Pearl Harbor: warning and decision*. Stanford, Calif.: Stanford University Press.
- Wolff, Christian. 1934. *The law of nations treated according to a scientific method*. Trans. Joseph H. Drake. Oxford: Clarendon Press; originally published 1749.
- Wollheim, Richard. 1958. Democracy. *Journal of the History of Ideas* 19: 225-42.
- Wollheim, Richard, ed. 1975. *John Stuart Mill, three essays*. Oxford: Clarendon Press.
- Wollheim, Richard. 1984. *The thread of life*. Cambridge: Cambridge University Press.
- World Commission on Environment and Development (WCED) (Gro Harlem Brundland, Chair). 1987a. *Our common future*. Oxford: Oxford University Press.
- World Commission on Environment and Development (WCED) (Gro Harlem Brundland, Chair). 1987b. Proposed legal principles for environmental protection and sustainable development. In WCED 1987a, pp. 348-51.
- Zeckhauser, Richard. 1975. Procedures for valuing lives. *Public Policy* 23: 419-64.
- Zeckhauser, Richard, and Elmer Shaefer. 1968. Public policy and normative economic theory. Pp. 27-102 in R. A. Bauer and K. J. Gergen, eds., *The study of policy formation*. Glencoe, Ill.: Free Press.

References

- Zimmerman, Michael J. 1985. Sharing responsibility. *American Philosophical Quarterly* 22: 115-22.
- Zimmerman, Michael J. 1987. Remote obligations. *American Philosophical Quarterly* 24: 199-206.

Name index

- Aaron, Henry J., 187, 192
Abel, Richard, 210
Ackerman, Bruce A., 160, 171
Adams, Robert Merrihew, 4, 16
Allais, Maurice, 292
Allen, H. M., Jr., 144
Allison, Lincoln, 59
Alternative Defence Commission, 304
American Law Institute, 212
Anderson, Elizabeth, 227
Anson, G. E. M., 10, 134
Arkin, William, 299
Arneson, Richard J., 125
Arrow, Kenneth J., 137, 142, 155, 204
Atiyah, P. S., 163, 164, 166, 175-8, 207,
214, 216, 224, 235
Atkinson, A. B., 188, 206, 241
Austin, J. L., 32
Austin, John, 12, 62
Avineri, Shlomo, 8
- Baker, Kurt, 81
Baker, John, 228
Ball, Desmond, 298
Banfield, Edward, 144
Barr, Nicholas, 204, 206, 229
Barry, Brian, ix, 7, 10, 13, 15, 20, 69, 119,
135-7, 147, 174, 206, 227, 245, 278, 283,
286, 305
Bator, F. M., 142
Baxter, R. R., 271
Bayes, T., 205-6
Beale, J. H., Jr., 298
Beitz, Charles, 73, 269, 277, 311
Benn, Stanley I., 125, 136, 138, 142, 143,
290, 296
- Bennett, Jonathan, 73
Bennett, Robert, 59
Bentham, Jeremy, 3, 5, 9-14, 16, 23-4, 62,
70
Bentzen, Lloyd, 232
Berger, Peter L., 233
Bertram, Chris, 206
Bessley, Tim, 206
Beveridge, William H., 186, 192, 230
Bickel, Alexander M., 270, 272
Bilmes, Jack M., 142
Bismarck, Otto von, 230
Black, Hugo, 224
Blaug, Marc, 229
Blechman, Barrie M., 289
Blough, Thomas, 161
Blumer, Martin, 233
Bracken, Paul, 298
Bradshaw, Jonathan, 229
Braithwaite, V. A., 233
Brandt, Richard B., 4, 7, 16, 55, 64, 65, 70,
137, 142, 143
Brandt, Willy, 47
Braybrooke, David, 245, 250
Brennan, Geoffrey, 58, 143, 180
Brewer, Daniel, 163, 209
Brierty, J. L., 271, 272
Brill, Alida, 147
Brittan, Samuel, 128
Brittan, Yvonne, 166, 167, 207, 218, 224
Broad, C. D., 296
Brody, Richard A., 289
Broome, John, 137, 174, 180, 207
Browning, E. K., 260
Brownlie, Ian, 271, 272
Brundland, Gro Harlem, 309, 311

Name index

- Buen, E. H., 194
 Burkhauser, Richard V., 161, 164, 167,
177, 224
 Burki, Shahid Javed, 245, 253
 Burns, Arthur Lee, 289, 305
- Cain, Bruce, 305
 Calabresi, Guido, 162, 174, 177, 178
 Campbell, Angus, 251
 Campbell, Tom, 206
 Cane, Peter, 163, 164, 166, 175-8, 207,
214, 216, 224
 Cannon, C., 146
 Cardozo, Benjamin, 269
 Castles, Francis G., 243
 Chambers, David L., 239
 Chase, Harold W., 272
 Coe, Richard D., 195-7
 Cohen, Joshua, 141
 Cohen, Marshall, 7, 311
 Cook, Philip J., 176
 Cooter, Robert, 19
 Corcoran, Mary E., 195-7
 Cordes, Joseph J., 171, 207, 215, 217
 Corfield, Peter, 166, 167, 207, 218, 224
 Corwin, Edward S., 160, 177, 213, 272
 Cottingham, John, 267, 283
 Cuwell, Frank, 305
 Crawford, James, 319
 Culyer, A. J., 164, 168, 224
- d'Abba, Peter, 233
 Daniels, Norman, 250
 Danziger, Sheldon, 24, 222
 Dasgupta, Partha, 75, 245, 255
 Davidson, Donald, 20, 50, 54, 104, 121, 122
 Davis, Nancy (Ann), 297
 Day, J. P., 163
 de-Shalit, Avner, 8
 Deacon, Alan, 229
 Dennett, Daniel, 6
 Derthick, Martha, 186-9
 Dickens, Charles, 5, 7
 Dockrill, Michael, 289
 Donagan, Alan, 66, 69, 70
 Douglas, Mary, 221, 255
 Dowding, Keith, 261
 Downs, Anthony, 143
 Doyal, Len, 250
 Drèze, Jean, 7, 250
 Dryzek, John, 141, 180, 183, 185, 229, 243,
294, 305
- Dubos, Rene, 306, 310
 Ducat, Craig R., 272
 Dulles, John Foster, 288
 Duncan, Greg J., 195-7
 Duncan, Otis Dudley, 251
 Dunleavy, Patrick, 141
 Dworkin, Gerald, 104, 126, 139
 Dworkin, Ronald M., 22, 23, 25, 119, 132,
134-7, 139, 146-7, 174
- Eads, George, 305
 Eckstein, Otto, 144
 Edwards, Alistair, 59
 Edwards, Meredith, 232, 239
 Elias, Robert, 167
 Elkin, Steve, 305
 Elliot, Robert, 170
 Ellis, Adrian, 250
 Ellsberg, Daniel, 288, 290, 292, 295
 Elster, Jon, ix, 15, 64, 128, 139-40, 148,
155, 256, 290, 292, 305
 Ely, John Hart, 59
 Engels, Friedrich, 253
 Enker, Arnold N., 167
 Epstein, Richard A., 160, 161, 207, 208,
209, 224
 Esping-Andersen, Gosta, 230
- Feinberg, Joel, 84, 89, 90, 98, 105, 125,
127, 162
 Feldstein, Martin, 161, 215, 221
 Fenn, Paul, 166, 167, 207, 218, 224
 Fiorina, Morris P., 145
 Fischhoff, Baruch, 290
 Fisher, Adrian S., 270, 271
 Fisher, Anthony C., 155
 Fishkin, James S., 30, 66
 Fitzmaurice, Deborah, ix, 227
 Fletcher, George P., 8-9, 55, 152
 Follesdal, Dagfin, 148, 290
 Frankena, William K., 283
 Frankfurt, Harry G., 138, 245, 251
 Freeden, Michael, 214
 Freedman, Lawrence, 299
 Frei, Daniel, 299
 French, Peter A., 35, 36
 Fried, Charles, 10, 136, 205
 Friedman, Milton, 6, 22, 228
 Friedman, Rose, 228
 Fuller, Lon, 220
 Fullinwider, Robert, 305

Name index

- Galanter, Marc, 258
 Ganßmann, Heiner, 261
 Garfinkel, Irwin, 239
 Gauthier, David, 289
 Genn, Hazel, 166, 167, 207, 218, 224
 Georgescu-Roegen, Nicholas, 168
 Gibbard, Allan, 137
 Gibson, Diane, 113, 183, 185, 229, 243
 Glaser, James M., 147
 Glendon, Michael, 319
 Glennester, Howard, 206
 Godwin, William, 267
 Goldfarb, Robert S., 171, 207, 215, 217,
221
 Goodin, Robert E., 6, 9, 12, 16, 19, 20, 30,
32-5, 41, 55, 65, 69, 70, 72, 76, 87, 90,
94, 109, 110, 119, 127, 128, 130, 135-8,
140, 142, 143, 145-8, 161, 170, 174, 178,
183, 185, 190, 200, 203, 204, 205, 207,
218, 228, 229, 233, 237, 241, 243, 245,
250, 258, 267, 269, 275, 278, 281, 285,
287, 290, 304, 313, 316
 Gorovitz, Samuel, 305
 Gough, Ian, 250
 Gradgrind, 7
 Graham, Daniel A., 176
 Grant, J. A. C., 213
 Gray, Colin S., 291, 302
 Gray, John, 135
 Green, Donald Philip, 147
 Gregory, Robin, 175
 Griffin, James, 162, 180
 Grigg, Charles M., 147
 Gross, B. M., 146
 Grotius, Hugo, 270, 277, 285

 Hagen, Ole, 292
 Haksar, Vinit, 59
 Halberstadt, Victor, 161, 164, 167, 177, 224
 Halberstam, David, 24
 Halévy, Elie, 12
 Hamblin, Charles, 292
 Hamermesh, Daniel S., 199
 Hamlin, Alan, 44
 Hampshire, Stuart, 3, 9, 10, 24
 Handler, Joel F., 228
 Hanson, Russell, 183, 185, 186, 229, 243
 Haq, Mahbub ul, 246, 253
 Hardin, Russell, ix, 3, 6, 7, 8, 16, 17, 18,
61, 147, 180, 300
 Hare, R. M., 4, 7, 9, 16, 17, 18, 65, 69,
268, 282, 283

 Harris, Donald A., 166, 167, 178, 207,
218, 224
 Harris, John, 22
 Harrod, Roy F., 60, 66, 76
 Harsanyi, John C., 64, 71, 136, 140-2,
162, 201
 Hart, Douglas M., 289
 Hart, H. L. A., 23, 25, 55, 65, 89, 108, 132,
134, 217, 267, 293, 294, 309
 Harvard Nuclear Study Group, 288, 289,
302
 Haveman, Robert H., 24, 161, 164, 167,
177, 183, 185, 222, 224, 229, 243
 Hayek, Friedrich A., 17, 73, 134, 190
 Hayter, Theresa, 47
 Head, John G., 136
 Heath, Anthony, 250
 Hegel, Georg W. F., 139-40
 Heil, John, 121
 Held, Virginia, 7, 35, 36, 42
 Hepple, B. A., 167, 176
 Heyd, David, 30, 66
 Hicks, John R., 19, 160, 171, 223
 Hicks, Norman, 246, 253
 Hill, Martha S., 195-7
 Hill, Thomas E., Jr., 143
 Hirsch, Fred, 147, 163, 200, 250, 254, 256
 Hitch, Charles J., 24
 Hochman, Harold M., 171, 221
 Hodgson, D. H., 18, 64, 71, 201
 Hoffman, Saul D., 195-7
 Hollis, Martin, 113, 138, 143, 227
 Holmes, Oliver Wendell, 120, 294
 Honoré, Tony, 18, 217, 277, 293, 294
 Hoole, Francis W., 290
 Hout, Michael, 147
 Hume, David, 64, 71, 127, 142, 201, 277
 Hunter, David, 59
 Hyland, Aanund, 148
 Hyman, H. H., 251

 Isherwood, Brian, 221, 255
 Ison, Terence G., 167

 Jackson, Frank, 33, 36
 Jackson, Kenneth, 254
 Jefferson, Thomas, 146
 Jervis, Robert, 293, 301
 Jevons, Stanley, 19, 121
 Johnson, Norman, 233
 Jones, Peter, 226
 Jones, Will, 305

Name index

- Kagan, Shelly, 65, 66, 69
 Kahneman, Daniel, 126, 127, 137, 153, 205, 290
 Kaldor, Nicholas, 19, 160, 168, 171, 223
 Kant, Immanuel, 18, 266-7
 Kaplow, Louis, 221
 Kavka, Gregory S., 290, 304, 305
 Keeny, Spurgeon M., Jr., 298
 Kennan, George F., 318
 Kennedy, John F., 24, 293
 Kennedy, Robert, 293
 Kevles, Dan, 305
 Kiewiet, D. Roderick, 144
 Kinder, Donald R., 144
 Kinninging, Andreas, 131
 Klawitter, Marielka M., 239
 Kleinig, John, 43, 174
 Knetch, Jack L., 175
 Kousser, J. Morgan, 305
 Kronman, Anthony, 210
 Kymlicka, Will, 8
- Lackey, Douglas P., 296, 300, 304, 305
 Lakatos, Irme, 162
 Lambert, Karel, 6
 Lancaster, Kelvin J., 20, 132, 169, 170
 Land, Hilary, 233
 Lau, R. R., 144
 Lave, C. A., 147
 Le Grand, Julian, ix, 180, 183, 185, 206, 229, 243
 Leader, Sheldon, 180, 227
 Lehrer, Keith, 180
 Levine, Henry D., 5
 Levmore, Saul, 227
 Lewis, David, 33
 Lindblom, C. E., 250
 Lindley, Richard, 131
 Lipsky, Michael, 194
 Little, L. M. D., 20, 141, 174
 Lively, Jack, 3
 Llewellyn, Karl, 120
 Lloyd-Bostock, Sally, 166, 167, 207, 218, 224
 Locke, John, 29, 124, 312
 Lomasky, Loren, 143
 Luce, R. Duncan, 290, 292, 295, 296
 Lukes, Steven, 9, 10, 12, 22
 Lyons, David, 16, 17, 76
- Maass, Arthur, 42
 Mabbott, J. D., 22, 62
- Macaulay, T. B., 23-4
 MacCallum, Gerald C., 58
 MacCormack, Geoffrey, 163
 MacCormick, Neil, 173
 MacIntyre, Alasdair, 51, 283
 Maciver, Robert M., 48
 Mack, Andrew, 305
 Mackie, J. L., 217
 MacLean, Douglas, 27
 Maclean, Mavis, 166, 167, 207, 218, 224
 MacRae, Duncan, Jr., 142
 Mansbridge, Jane J., 8, 142
 March, James G., 5, 147
 Marglin, Stephen A., 129, 144
 Margolis, Howard, 142, 180
 Marmor, Ted, 206
 Marshall, Alfred, 19, 49
 Marx, Karl, 253
 Mathews, Jessica Tuchman, 310
 May, Ernest, 146
 Mayo, H. B., 134
 McCarthy, Peter, 27
 McCloskey, Herbert A., 147
 McDaniels, Tim, 175
 McGraw, Kathleen M., 126
 McKean, Roland N., 24
 McKiney, Michael, 43
 Meade, James E., 279
 Melamed, A. Douglas, 162, 174, 177, 178
 Mellema, Gregory, 90
 Mendus, Susan, 131
 Menzies, Peter, 305
 Michelman, Frank J., 160, 171, 207, 215, 217, 224
 Mill, James, 5, 12, 23-4, 128, 129
 Mill, John Stuart, 9, 12, 13, 62, 120, 123, 125, 127, 128, 129, 135, 141
 Miller, David, ix, 59, 131, 206, 279, 286
 Mills, C. Wright, 49
 Mises, Ludwig von, 120
 Mishan, E. J., 162, 175, 179
 Mondale, Walter F., 232
 Montreal Protocol, 318-19
 Moore, G. E., 2
 Morgan, James N., 195-7
 Morgenbesser, Sidney, 2
 Muller, Ned, 206
 Murray, Charles A., 229, 233
 Musgrave, Alan, 162
 Musgrave, Richard A., 136, 142
 Myrdal, Gunnar, 147

Name index

- Nagel, Thomas, 7, 8, 10, 72, 269, 282
 Narveson, Jan, 301
 Neuhaus, Richard J., 233
 Neutze, Max, 59
 Nisbett, Richard E., 51
 Nonet, Philippe, 167, 192
 Norman, Richard, 123
 Novack, Michael, 233
 Nozick, Robert, 23, 25, 34, 43, 44, 54, 69,
90, 106, 128, 155, 160, 161, 162, 168,
173, 178, 201, 208, 209, 217
- Offe, Claus, 203, 206, 230, 241, 243
 Okin, Susan Moller, 8
 Okun, Arthur, 229
 Olsen, Johan P., 5
 O'Neill, Onora, ix-x, 174, 227
 Oppenheim, L., 270-2, 285
 Organisation for Economic Co-operation
 and Development (OECD), 233
- Page, Talbot, 200, 204
 Pahl, Jan, 232, 239
 Palme, Joakim, 230
 Panofsky, Wolfgang K. H., 298
 Pareto, Vilfredo, 168
 Parfit, Derek, 14, 72, 73, 128, 144, 268
 Partington, Martin, 194
 Paskins, Barrie, 289
 Payne, Keith B., 291
 Pears, David, 153
 Pechman, Joseph A., 187, 192
 Perlman, Moris, 227
 Perot, H. Ross, 218
 Peters, B. Guy, 220
 Peters, R. S., 48, 49
 Pettit, Philip, x, 9, 30, 35, 40, 44, 59, 72,
87, 113, 180, 281, 285, 313
 Piachaud, David, 232, 239, 245
 Pigou, A. C., 19, 20, 70, 144, 222
 Plamenatz, John, 12, 142
 Plant, Raymond, 59, 225
 Plotnick, Robert, 24, 222
 Price, Huw, 305
 Pringle, Peter, 299
 Prior, A. N., 293
 Prosser, William, 90, 160, 161, 165, 166,
176, 177, 212, 293, 294
 Prothro, James W., 147
 Pufendorf, Samuel, 90
- Quester, George, 305
- Raiffa, Howard, 54, 290, 292, 295, 296
 Raphael, D. D., 293
 Rappoport, Peter, 19
 Rawls, John, 19, 23, 62, 64, 65, 70, 277
 Ray, Debraj, 246, 255
 Raz, Joseph, 119, 134
 Rees, John, 3
 Reeve, Andrew, x, 9, 59, 119, 126, 131
 Reeve, S., 176
 Regan, Donald H., 18, 60, 66
 Reich, Charles A., 228
 Ribicoff, Abraham, 232
 Richard I (king of England), 255
 Riley, Patrick, 136
 Robbins, Lionell, 19, 121
 Roberts, K. W. S., 142, 143
 Roosevelt, Franklin D., 189
 Rose, Richard, 220
 Rosenberg, Gerald M., 272
 Rosenthal, Joel, 319
 Ross, W. D., 6
 Rothenberg, Jerome, 138, 139
 Rousseau, Jean-Jacques, 136, 244
 Runciman, W. G., 142, 251
 Ryle, Gilbert, 50
- Sagan, Carl, 301
 Sagan, Scott D., 288, 289, 302
 Sagoff, Mark, 170
 St. Leonards, Lord Chancellor, 194
 Sandel, Michael J., 8
 Sax, Joseph L., 160
 Saylor, Taylor Reese, 47
 Scanlon, T. M., 7, 123, 311
 Scheffler, Samuel, 8, 73
 Schlesinger, Arthur M., Jr., 288
 Schoeck, Helmut, 255
 Schofield, Norman, 305
 Schuck, Peter H., 279
 Scitovsky, Tibor, 223
 Searn, David O., 144
 Sejersted, Francis, 148
 Self, Peter, 142
 Sen, Amartya, 7, 8, 13, 15, 19, 21, 48, 58,
61, 62, 72, 132, 133, 137, 138, 142, 143,
162, 163, 168, 169, 174, 200, 204, 249,
250
 Shacknové, Andrew E., 285
 Shaefer, Elmer, 162
 Shklar, Judith N., 10, 146
 Shubik, Martin, 250
 Shue, Henry, 42, 73, 269, 281, 283, 305

Name index

- Sidgwick, Henry, 7, 16, 21, 49, 61-2, 70,
161, 164, 173, 268, 271, 272, 277, 283
- Simmons, A. John, 311
- Sinden, J. A., 175
- Singer, Peter, ix, 7, 66
- Skinner, B. F., 143
- Slovic, Paul, 127, 153, 205, 290
- Smart, J. J. C., 6, 12, 16, 17, 133, 305
- Smith, Adam, 121, 127, 142, 245
- Sniderman, Paul M., 147
- Snyder, Jack, 293, 301, 305
- Sohn, Louis B., 271
- Stapleton, Jane, 214
- Steiner, Hillel, 59, 228, 287
- Stephen, James Fitzjames, 135
- Stewart, James, 246
- Stewart, Frances, 246, 253
- Stoics, 15
- Stone, Deborah, 235
- Stone, Peter B., 306
- Streeten, Paul, 246, 253
- Stretton, Hugh, 59
- Suárez, Francisco, 270, 271, 272
- Sunstein, Cass R., 141, 148
- Suppes, Patrick, 54
- Sylvan, Richard, 59, 305
- Taussig, Michael K., 187, 192
- Taylor, C. C. W., 123
- Taylor, Charles, 162
- Taylor, Michael, 40, 59, 148, 317
- tenBroek, Jacobus, 224, 232, 235
- Tetlock, Philip E., 147
- Thompson, Dennis F., 10
- Thomson, Judith Jarvis, 104, 161, 175, 178
- Titnuss, R. M., 190, 229, 235, 237
- Tobin, James, 221, 260
- Townsend, Peter, 245-7
- Trachtenberg, Marc, 289, 299, 305
- Tribe, Laurence H., 160, 162, 171
- Tulloch, Gordon, 171, 173, 178, 180, 221
- Tussman, Joseph, 235
- Tutu, Desmond, 48
- Tversky, Amos, 126, 127, 137, 153, 175,
205, 290, 292
- Tyler, Tom R., 144
- United Nations Conference on the
Human Environment, 306, 310, 311,
317
- United States Department of Health,
Education, and Welfare (DHEW), 185,
193, 195, 230
- United States Senate, 146
- Urmason, J. O., 30, 62, 66, 150
- van der Veen, Robert, 228
- van Guinsteren, Herman, 148
- Van Parijs, Philippe, 228, 229, 243, 261
- Vattel, Emerich de, 270, 271, 277, 285
- Vienna Convention, 318-19
- Wade, J. W., 160, 163, 165, 166, 176, 177
- Waldron, Jeremy, 21, 27, 44, 59, 113, 225,
253
- Walsh, Clifford, 136
- Walzer, Michael, 10, 260, 277, 311
- Ward, Barbara, 306, 310
- Ward, Hugh, 40, 59, 141, 317
- Ware, Alan, 261
- Wasserstrom, Richard, 33, 283, 286
- Watson, Gary, 138
- Watt, A. J., 123
- Weale, Albert, 59, 227
- Weaver, R. Kent, 126
- Weber, Max, 10
- Weinrib, E. J., 43
- Weitzman, M. L., 260
- Wickstead, Philip H., 168
- Wiggins, David, 245
- Wildavsky, Aaron, 291
- Williams, Bernard, 8, 10, 21, 23, 54, 61,
62, 68, 69, 121, 143, 153, 162, 173, 201,
267, 268, 282, 289, 296
- Williams, Glenville, 90, 167, 176
- Williston, S., 162, 166
- Wilson, James Q., 144
- Wilson, Timothy DeCamp, 51
- Winter, David, 183, 185, 229, 243
- Wirth, David A., 307
- Wohlstetter, Roberta, 288
- Wolff, Christian, 270, 271, 272, 277, 283,
285
- Wollheim, Richard, 134, 155, 173
- World Commission on Environment and
Development (WCED), 309, 311
- Zeckhauser, Richard, 162, 176, 290
- Zimmerman, Michael J., 32, 90
- Zinnes, Dina A., 290

Utilitarianism, the great reforming philosophy of the nineteenth century, has today acquired a reputation for being a crassly calculating, impersonal philosophy unfit to serve as a guide for moral conduct. Yet what may disqualify utilitarianism as a personal philosophy makes it an eminently suitable guide for public officials in the pursuit of their professional responsibilities.

Robert E. Goodin, a philosopher with many books on political theory, public policy, and applied ethics to his credit, defends utilitarianism against its critics and shows how it can be applied most effectively over a wide range of public policies. In discussions of such issues as paternalism, social welfare policy, international ethics, nuclear armaments, and international responses to the environment crisis, he demonstrates what a flexible tool his brand of utilitarianism can be in confronting the dilemmas of public policy in the real world.

Written in a lucid, nontechnical style, these essays will interest a large cross section of the academic community concerned with, and teaching courses on, public policy, whether they be in departments of philosophy, political science, law, or economics.

Cover design by Dennis M. Arnold

CAMBRIDGE
UNIVERSITY PRESS

ISBN 0-521-46806-X



9 780521 468060