

American Economic Association

Taking Ethics Seriously: Economics and Contemporary Moral Philosophy

Author(s): Daniel M. Hausman and Michael S. McPherson

Source: *Journal of Economic Literature*, Vol. 31, No. 2 (Jun., 1993), pp. 671-731

Published by: American Economic Association

Stable URL: <http://www.jstor.org/stable/2728513>

Accessed: 06/03/2009 13:02

Your use of the JSTOR archive indicates your acceptance of JSTOR's Terms and Conditions of Use, available at <http://www.jstor.org/page/info/about/policies/terms.jsp>. JSTOR's Terms and Conditions of Use provides, in part, that unless you have obtained prior permission, you may not download an entire issue of a journal or multiple copies of articles, and you may use content in the JSTOR archive only for your personal, non-commercial use.

Please contact the publisher regarding any further use of this work. Publisher contact information may be obtained at <http://www.jstor.org/action/showPublisher?publisherCode=aea>.

Each copy of any part of a JSTOR transmission must contain the same copyright notice that appears on the screen or printed page of such transmission.

JSTOR is a not-for-profit organization founded in 1995 to build trusted digital archives for scholarship. We work with the scholarly community to preserve their work and the materials they rely upon, and to build a common research platform that promotes the discovery and use of these resources. For more information about JSTOR, please contact support@jstor.org.



American Economic Association is collaborating with JSTOR to digitize, preserve and extend access to *Journal of Economic Literature*.

<http://www.jstor.org>



Taking Ethics Seriously: Economics and Contemporary Moral Philosophy

By DANIEL M. HAUSMAN
University of Wisconsin-Madison

and

MICHAEL S. MCPHERSON
Williams College

We are especially indebted to John Roemer for proposing this project and for offering extensive criticisms of earlier drafts. Richard Arneson, Henry Bruton, Eric Kramer, Philippe Mongin, Morty Shapiro, Amartya Sen, Julius Sensat, Hamish Stewart, Gordon Winston, and anonymous referees corrected many mistakes and led us to additional sources. Anne McEachern and John McEachern provided valuable research assistance. Hausman gratefully acknowledges the support of a Vilas Associate award from the University of Wisconsin-Madison.

TO BE A GOOD PERSON, one must take ethics seriously. But can the same be said about being a good *economist*? Does morality matter to economic analysis? It would be idle to suggest that economists should spend a large part of their time studying moral philosophy, and moral sophistication may be more important in some areas of economics, such as political economy, than in others, such as econometrics. Yet we would argue that a moderate dose of reflection on the relations between economics and morality could help many economists work more effectively. We should emphasize our concern to show the relevance of moral thinking to the conduct of *both* positive

and normative economics. Not only do moral principles bear on issues concerning evaluation and policy, but they also influence the questions positive economists ask and the answers they find plausible.

The present essay surveys recent work by economists and moral philosophers that borders the two disciplines. It aims to provide an overview for a broad audience of working economists and to suggest possibilities for research to those with specific interests in “ethics and economics.” (Book length treatments with similar goals include Allen Buchanan 1985 and Alan Hamlin 1986; see also Sen 1987a.) After the introductory and moti-



vational remarks of Section 1, Section 2 will develop the relations between morality and rationality and the problems of modeling moral commitments. Section 3, the longest section in the essay, presents a number of different perspectives for the ethical evaluation of economic processes and outcomes. Section 4 is concerned with the moral relevance and presuppositions of social choice theory and game theory.

1. *Economics and Moral Theory*

Before beginning, a few abstract words on the relations between economics and ethics are in order. Those who deny that economics and ethics having anything in particular to do with one another rarely deny that economics is relevant to policy. They maintain instead that the relevance of economics is purely technical. It provides causal knowledge of the consequences of policies to enable policy makers to choose effective means toward their ends. Ethics determines the ends, and economics determines the means. They are both crucial to policy, but they have nothing much to do with one another.

The simple picture of the economist who provides value-free technical information to the decision maker is at best a useful caricature. It fits the activity of an economist who calculates the revenue losses from a proposed tax reduction, but it does not fit the activity of an economist who is asked for advice. For the political process never formulates all the relevant goals and constraints explicitly or specifies how to weight them (Fritz Machlup 1969, pp. 119–24).¹ If President Carter had asked Milton Friedman or Paul Samuelson, “What is the ‘best’ way to reduce inflation?” neither economist would have

¹ We are indebted to Samuel Weston (forthcoming) for this reference.

had a well-defined technical problem until he figured out what other objectives Carter had and what their relative weights were. At some point they would almost certainly have had to rely on some of their own values, too. The same issue arises in selecting problems to study: economists who refuse to “dirty their hands” with ethical matters will not know what technical problems to investigate.

Positive and normative questions may interact in the conduct of research as well as in deciding what problems to study or what advice to give. As some moral views “fit” familiar patterns of economic modeling more easily than others, economists may focus on the moral views that best mesh with their theoretical preconceptions, or distort other moral views to make them fit. Better understanding of moral views that are hard to link to familiar patterns of economic reasoning, such as those that emphasize rights or needs, could influence the development of economics, as theoretical categories are modified to provide other ways of organizing experience. Working out the implications of different moral views may also suggest novel economic problems or theoretical formulations.² The traffic need not all go from philosophy to economics. Theoretical results in economics such as Kenneth Arrow’s impossibility theorem or Sen’s liberal paradox force moral theorists to rethink their principles. Economic theorists have developed many analytical tools—in game theory, social choice theory, and the measurement of welfare, among others—that have challenged moral philosophers to articulate their views better and have

² It may be that development of “property rights economics” has been influenced in this way by the libertarian orientation of some its practitioners, just as Marxian commitments have influenced the development of theories of exploitation by John Roemer and others.

helped them to do so.³ Intellectual progress is either moral philosophy or economic theory may stimulate workers in the other discipline.

1.1 *Why Should Economists Be Interested in Moral Questions?*

Economists should care about moral questions for at least the following four reasons:

1. The morality of economic agents influences their behavior and hence influences economic outcomes. Moreover, economists' own moral views may influence the morality and the behavior of others in both intended and unintended ways. Because economists are interested in the outcomes, they must be interested in morality.
2. Standard welfare economics rests on strong and contestable moral presuppositions. To assess and to develop welfare economics thus requires attention to morality.
3. The conclusions of economics must be linked to the moral commitments that drive public policy. To understand how economics bears on policy thus requires that one understand these moral commitments, which in turn requires attention to morality.
4. Positive and normative economics are frequently intermingled. To understand the moral relevance of positive economics requires an understanding of the moral principles that determine this relevance.

³There are concerns among some contemporary philosophers that the influence of economic theory on the development of modern moral philosophy has been all too great, leading to a neglect both of social history and of the history of philosophy as sources of understanding in this realm. See Michael Walzer (1983) and Charles Taylor (1989) for sustained attempts to carve out other paths.

1.1.1 *The Economic Importance of Morality.* Morality influences economic outcomes. The morality accepted by economic agents influences how they behave. It is among the causal factors that economists need to study. For example, as Fred Hirsch (1976) and others have noted, economic efficiency and economic policy depend crucially upon ethical values which may be undermined by the development of market economies (Karl Polanyi 1944; Albert Hirschman 1985; Sen 1977a; Arrow 1974). Without honesty, trust, and goodwill, economic life would grind to a halt. Insofar as economists seek to explain economic outcomes and to promote economic efficiency, they need to be concerned with the nature and sources of useful or harmful moral virtues or vices. Can economically useful virtues be sustained by rational self-interest, or does the success of economies depend crucially upon non-self-interested moral commitments? The role of moral norms in accounting for economic outcomes has received increasing attention in the areas of labor economics, economics of organization, and economic development. These developments are surveyed in Section 2.2 below.

For a striking example of the economic importance of moral norms, consider the work of Richard Titmuss (1971) concerning systems for acquiring and distributing blood for transfusions. Titmuss claims to show that voluntary blood donation systems, such as Great Britain's, are more efficient than are commercial systems, and he argues that they embody and cultivate altruistic communitarian values. In a thoughtful and lengthy review, Arrow endorses much of Titmuss' case. In particular, Arrow underscores the serious drawbacks of commercial provision when the testing of blood for infectious agents is imperfect. Altruistic donors have no incentive to lie about whether their blood is safe, but commercial donors plainly do.

Arrow does, however, question Titmuss' ambitious claim that the commercial option for donors in the United States undermines the willingness of U.S. citizens to give blood altruistically (1972, esp. pp. 350–51). In Arrow's view markets increase freedom and enable us not ". . . to use up recklessly the scarce resources of altruistic motivation" (1972, p. 355). Arrow can make little sense of Titmuss' claim that markets "deprive men of their freedom to choose to give or not to give" (1972, p. 350). But as Peter Singer (1973, 1977) and Hirschman (1985) have noted, it can be misleading to regard altruism as a scarce resource, for, like a skill, its quantity can be increased by its use. And although Titmuss' language is puzzling, commercialization can discourage giving. To give blood, when it cannot be purchased, is to give something that is literally priceless, something which the donors themselves describe as "the gift of life." To give blood when it can be purchased is to give the equivalent of \$50 (Singer 1977). Whether or not these attitudes are rationally justified, their potential economic effects are evident, and they can only be understood if economists are willing to study some ethics.

Even if economists should attend to the "sociological" influence of moral commitments on economic outcomes, it does not follow that they need to worry about the substantive "truth" or validity of those moral commitments. But it seems to us that economists may be able to advance their work by appraising people's moral behavior and ideas as well by tracing their causal consequences. It is natural, illuminating, and virtually unavoidable for economists to inquire whether people's observed (or alleged) moral commitments "make sense." Part of Arrow's discomfort with the Titmuss argument was that there seemed to him to be no good *reason* why people's willingness to donate blood should be influenced by the existence of a market. If

an alleged preference or moral attitude is irrational, its existence and stability are cast into doubt. Descriptions of moral motivations and judgments are likely to be evaluations too, and are not always less useful for that.

Furthermore, as the examples suggest, the moral convictions of economic agents, unlike causal factors such as scarcities, can be influenced by the way in which they are analyzed and described by economists. Saying that human behavior can be modeled as if it were entirely self-interested (as the Chicago economists in particular do) or defining rationality as self-interest (as in Robert Frank 1988, p. 2) legitimizes and fosters self-interested behavior. In experiments on human cooperation, economics and business students are uniquely uncooperative (Gerald Marwell and Ruth Ames 1981). Learning economics, it seems, may make people more selfish.⁴ A causal factor that depends upon the way in which it is described by economists is a tricky thing to study.

The facts that economists can rarely describe moral commitments without evaluating them and that economists affect what they see by how they describe it, provide reason for them to think about not just the morality in fact accepted in the society they study but about the morality they think should be accepted. Moral reflection is important for positive economics, not just normative economics. The moral principles accepted by positive economists matter empirically because they are economic agents as well as theorists. Their theorizing reflects their moral commitments and influences the actions and moral commitments of others.

1.1.2 *Morality and Standard Welfare*

⁴Or at least more calculating. It has been suggested that economists are more "selfish" in laboratory "prisoners' dilemma" experiments because they grasp the situation better. Their unwillingness to cooperate could also reflect previous self-selection.

Economics. Welfare economists have sometimes thought that their techniques are virtually morally neutral, because they rest on innocuous and uncontroversial moral premises such as the Pareto principle. But the standard definition of a social optimum compares social alternatives exclusively in terms of the goodness of their outcomes (rather than the rightness of their procedures) and identifies the goodness of outcomes with satisfaction of individual preferences. These commitments to value only outcomes and to measure outcomes only in terms of individual utilities are neither neutral nor uncontroversial (see Sen 1979b; and Sen and Bernard Williams 1982).

Economists do not of course limit their contributions to policy discussions to identifying Pareto improvements. Familiar practices such as cost-benefit analysis and routine preferences for competitive markets cannot be defended in purely Paretian terms. In fact, developments within welfare economics itself over the last fifty years, including the theory of the second-best (Richard Lipsey and Kelvin Lancaster 1956) and the criticism of the so-called “compensation principle” (Samuelson 1950a) have made clear that no single all-purpose defense of these familiar practices and judgments is available (see Section 3.2.1 below). Yet the strategy of retreating into silence whenever the search for pure Pareto improvements comes up empty is not a plausible one. Because normative economists must make further evaluations of economic processes and outcomes, they cannot avoid moral assessment of the techniques and principles they rely on. It may be possible to reinforce standard policy conclusions by articulating a broader range of values, as in arguing for competition on grounds of its contributions to personal liberty as well as to efficiency.⁵ Or

⁵ Such arguments are stressed by authors as different as Friedman (1962), John Rawls (1971), and Robert Nozick (1974) as well as in the tradition of classical

it may be possible to improve on familiar techniques and to reach somewhat different policy conclusions. For example, one can correct for the possible distributional inequities of an unconstrained cost-benefit analysis (Anthony Atkinson 1970; Arnold Harberger 1978). Controversial moral commitments are unavoidable when the arguments are deepened in these ways.

Much applied welfare economics factors normative questions into questions of efficiency and of equity. This factoring is motivated by two things.⁶ First, theoretical economics provides apparently precise and tractable notions of efficiency, while some notions of equity are vague, and most have little connection to the terms in which economists theorize. Second, the values that support greater efficiency have seemed less controversial than those supporting greater equity. So problems of efficiency are seen as manageable and uncontroversial, while problems of equity are seen as alien, controversial, and theoretically intractable. An unsurprising consequence is that standard welfare economics focuses on efficiency.

This circumstance already presupposes a great deal of moral theory. What is the moral basis of a concern with efficiency? Is it really less controversial than the moral commitments that lie behind notions of equity? What do economists presuppose when they factor questions of economic welfare into questions of efficiency and questions of equity? Can these moral presuppositions be sustained? Think about the extreme example of applying cost-benefit analysis to decide whether to build nuclear power

liberalism, for instance John Stuart Mill (1871, book 11, ch. 1) and Frank Knight (1935).

⁶ For an excellent early discussion and critique of the alternative ways in which economists employ and defend the partitioning of normative questions into equity and efficiency components, see Paul Streeten (1953). See also Ian Little (1957).

plants. How should the interests of future people be reflected? What (if any) discount rate is morally appropriate to use? What is the right way of dealing with the tremendous uncertainties of the distant future? (Partha Dasgupta 1982; Isaac Levi 1980; Robert Lind et al. 1982; Douglas Maclean and Peter Brown 1983; Dennis Mueller 1974; Derek Parfit 1984, pp. 480–86; Peter Warr and Brian Wright 1981.)

1.1.3 *Making Economics Relevant to Policy.* When politicians and noneconomists think about problems of welfare, they employ concepts that do not easily translate into the language of standard economic theory. The Reagan administration spoke about preserving a “safety net.” Panhandlers speak to us of their needs, not their preferences. Notions of fairness, opportunity, freedom, and rights are arguably of more importance in policy making than are concerns about moving individuals up their given preference rankings. Insofar as economists want to assist in the formulation of policy toward such ends, they must link economic theory to such concerns. Doing so requires understanding what these vague things are that people apparently value so much.

Consider, for example, the question of whether government should provide food or health care in kind, or whether it should instead provide a cash supplement to those in need. Most economists would offer a simple argument why the government should supply cash:⁷ Construct an indifference map for an individual with quantity of health care on one axis and some composite consumption

commodity on the other. Increased income raises the budget line. If the income increase is as large as the value of the in-kind health care provision, then the individual can purchase what the government would have provided if he or she prefers, or the individual can instead consume more of something else. The cash supplement makes individuals at least as well off as does the in-kind provision. Cash supplements are not only Pareto-superior to in-kind provision, but they also serve freedom by increasing the range of individual choice. And lower administrative costs would permit lower taxes or more health care.

Although this analysis brings in the value of freedom, it works mainly within the terms set by orthodox economic theory. There is no mention of needs, of the presuppositions of individual dignity, of opportunity, of rights, or of fairness. There is no concern with the normative commitments that support taxing individuals to provide such benefits (Steven Kelman 1986; Harold Hochman and James Rogers 1969). Not all the arguments in favor of in-kind provision that might follow from these points are paternalistic, but serious attention to questions of justified paternalism are begged from the beginning by equating recipients' welfare with the satisfaction of their preferences (which may be short-sighted or ill-informed).⁸ So even if the simple argument were valid, it would be inconclusive. How much weight it ought to have in policy making cannot be considered without a much broader concern with ethical questions.

⁷ For criticisms of this argument, see Lester Thurow (1977) and Arthur Okun (1975). This argument is faulty, for it leaves out of consideration the preferences of the donors. Asymmetric information also gives rise to serious problems (Charles Blackorby and David Donaldson 1988). If cash rather than medical care is provided to those in poor health, people will have an incentive to pretend to be sick.

⁸ Notice that one might acknowledge the possibility that persons' preferences are flawed indications of their welfare and still argue on grounds of liberty against paternalistic intervention. Mill's classic critique of paternalism in *On Liberty* does not assume that individuals are always the best judge of their own good. For assessments of apparently paternalistic interventions, see Gerald Dworkin (1971, 1988), and Joel Feinberg (1986).

For another illustration, think about the apparently irrational way in which most societies allocate resources to the preservation of life and health. If the money spent on neonatal intensive care units were devoted instead to better health care for pregnant women and less precarious newborns, infant mortality could be reduced significantly. If the heroic efforts to rescue victims of coal-mine cave-ins were devoted instead to mundane mine safety, fewer miners would die. Hundreds of thousands of dollars were spent to rescue Baby Jessica, who fell into an abandoned well, money which could instead have been used to prevent dozens of avoidable deaths of other little children. According to some moral theories, such as utilitarianism, these choices are apparently as indefensible as most economists would find them. But what sort of a social world would we live in, if our humanitarianism were always under the control of rational calculation, if we said to Baby Jessica's mother or to the families of trapped coal miners that the victims could be saved, but it would cost too much to save them? Does rational beneficence (a virtue according to almost all moral theories) demand that we "humanely" pump down poison gas to shorten the suffering? While these examples dramatize life-and-death dilemmas, a great many economic choices require weighing numbers of deaths against other values, or weighing the deaths of some people against the deaths of others (perhaps in future generations). Economists who work on such matters need to understand the strongly felt moral values that govern such choices and to form a sensible appraisal of these values.

1.1.4 *On the Intermixture of Positive and Normative Economics.* Many economists would maintain that questions of positive economics—of how economies actually function, of what the conse-

quences of policies will be—are entirely distinct from questions of normative economics—of what values economic arrangements should promote and of what policies should be adopted. Philosophers have, with good reason, questioned whether a sharp distinction can be maintained (John McDowell 1978; Hilary Putnam 1981, ch. 6; Richard Rorty 1982, ch. 11). But even if philosophers are wrong and there is in principle a sharp distinction to be drawn, the fact is that one often finds positive and normative concerns intermingled in economics.

For example, Paul Samuelson's influential essay, "An Exact Consumption-Loan Model of Capital and Interest" begins with a positive question concerning the effects on interest rates of the desire to save for retirement, but shifts midway to the normatively charged and highly theoretical question of the efficiency of competitive markets in infinite economies. And the two major initial comments on Samuelson's paper by William Meckling (1960) and Abba Lerner (1959) stem respectively from a concern to defend competitive markets and a concern to defend the welfare state. Normative issues have continued to play an important role in the development and assessment of overlapping-generations models (Hausman 1992, ch. 7).

Given shared moral commitments, theories that appear to be purely positive can have sharp normative consequences. For example in "The Economic Consequences of Cognitive Dissonance" (1982), George Akerlof and William Dickens offer a positive model of what will happen if individuals are able to adjust their beliefs to reduce cognitive dissonance. But the possibility of such adjustments casts doubt upon the identification of well-being with the satisfaction of preferences and consequently has immediate ethical consequences. If in order to reduce cognitive dissonance

individuals adjust their beliefs concerning, for example, the hazards of their work, there is a stronger case to be made for government safety regulations. One can understand how positive theorizing about cognitive dissonance or other features of individual decision making can have so much normative importance only if one appreciates the ethical presuppositions from which its normative punch derives.

The same point can be put in a different way. Despite (or perhaps because of) their insistence that they need to make no controversial moral commitments, economists often speak on policy questions with an appearance of great moral *authority*. What are the moral foundations of that apparent authority? What sort of moral principles do most economists accept? Why do they seem so uncontroversial?

So there are a number of good reasons why economists need to think about morality. For only by thinking about morality can they understand why positive theory can be of so much moral importance, what values govern actual policy making, what moral presuppositions govern welfare economics, and what moral factors influence economic outcomes.

1.2 *What Does Moral Philosophy Have to Contribute?*⁹

That economists face ethical questions does not imply that they will find answers in the scholarly work of moral philosophers. Because economics is a specialized discipline, with its own tools and outlook, one suspects economists will have to work out many of the answers on their own. **Still, economists may be pleasantly surprised by how much work in contemporary philosophy is useful to them.** Apart from systematic theorizing, such as new elaborations and defenses of utilitarian or contractualist theories, there is a wealth of narrowly focused

work that may be of use to economists. Some contemporary moral philosophy will be familiar and usable precisely because it builds with tools philosophers have borrowed and adapted from economists. Reading some moral philosophy may also help exorcise unjustifiable skeptical doubts (to which economists have been prone) about the possibility of rational argument concerning values. Finally, the literature of moral philosophy will be useful as a warning against reducing the many dimensions of moral appraisal to one or two and as a reminder of how many different dimensions of moral appraisal there are.

The last point is worth elaborating. **Philosophers have attended to many complexities of moral thought to which economists have paid little attention.** Unless moral philosophy is to be seen as a purely a priori discipline, which is neither a plausible view nor a popular one among philosophers, part of its task must be to make sense of, and bring some order to, everyday moral experience.⁹ On some issues (such as abortion) people have strong convictions. On others (such as animal rights), they may have mostly doubts. On some moral questions (should the homeless be allowed to starve?) people reach agreement easily, while on others controversy rages. Moral judgments are made about a variety of kinds of "ob-

⁹ This is not to say that moral philosophers must be uncritical of "commonsense morality." There is some analogy here to natural science: scientists must start from everyday observations in building up a scientific picture of the world, but their conclusions may wind up being critical of some commonsense beliefs about the world. They could not, however, be critical of all everyday observations at once: everyday observation is the raw material of science. Similarly, everyday moral judgments are the raw material of moral philosophy, which moral theory has to connect up with, but does not in every case have to endorse. This is a rough statement of the view of moral justification Rawls has called "reflective equilibrium" (1951, 1971). It is helpfully discussed in Norman Daniels (1979), and Henry Sidgwick's classic discussion of moral intuitions (1901) is still of value.

jects”: we may pronounce a person, an outcome, an action, an emotion “good” or “bad” in a moral sense. The complexity and fragmentation of actual moral judgments poses an important challenge to theory, one that is very familiar to economists in other contexts. There is an apparently inescapable trade-off between simplicity, generality, and theoretical precision on the one hand, and plausibility and recognition of complexity and “messiness” on the other.

This challenge is felt by economists as well as moral philosophers. But philosophers can perform a real service for economists to the extent that their work calls attention to the complications and multiple dimensions of moral judgment. The service is in part the negative one of discouraging economists from premature or overly sweeping generalization. But it includes as well the positive service of presenting economists with a variety of normative phenomena to which they may contribute their modeling skills. Indeed, it seems reasonable to suppose that both economists and philosophers may contribute as much to improving our understanding of moral problems by analyzing limited pieces of the puzzle as by advancing more general theories of morality. Much of the work we shall discuss is of the former kind.

2. *Morality and Rationality*

In their everyday usages, both “moral” and “rational” are normative and expressive terms: to characterize a choice as irrational or immoral is to condemn it, and not simply to describe it (Allan Gibbard 1990). Because modern economics is heavily committed to a specific characterization of rational conduct, it is committed to certain norms of rational individual choice. How are these norms related to moral norms, and how do moral norms influence what people do?

Attempting to bring moral notions “into the picture” of the economic actor clearly raises a number of questions.

In sorting all this out, it may be useful to follow two separate lines of inquiry. One involves attempts by economists to take account empirically of the influence of moral norms on people’s behavior. This approach starts from the apparent fact that people’s economic behavior is influenced by their moral beliefs, and it looks to see what impact these beliefs have on economic outcomes. At least to begin with, one can pursue this approach in a purely “positive” light—asking not whether these moral beliefs are justified, but simply whether they are behaviorally effective. Rather quickly, however, such work turns to the question of whether the observed moral commitments are economically beneficial or pernicious, and at that point the normative and positive dimensions intertwine. Work on these lines tends to use rough and ready conceptions of economic rationality and of morality—the aim is to find categories that help sort out empirical phenomena.

The other line of inquiry is more self-consciously conceptual and theoretical. Viewed abstractly, what is the relation between a morality and a preference relation? Can one or should one offer a conception of economic rationality that makes sense of a distinction between acting morally and maximizing utility? If there are cases where the two can be said to diverge, does this imply that it is sometimes irrational to act morally or immoral to act rationally?

It will be useful before pursuing these two lines of inquiry to review briefly the main features of the standard view of economic rationality.

2.1 *Economic Rationality*

Economics portrays agents as choosing rationally. Many generalizations concerning how people do in fact choose are

also claims about how agents *ought rationally to choose*. This fact distinguishes economics from the natural sciences, whose particles do not choose at all, and whose theories have no similar normative dimension. Economists take choice as arising from constraints, preferences, and expectations or beliefs. Economists typically take preferences as givens and not themselves subject to rational appraisal. But choices and sets of beliefs and preferences may be rational or irrational. Choice is rational when it is determined by a rational set of beliefs and preferences. The rationality of sets of preferences and beliefs is defined within utility theory.

2.1.1 *Certainty and Ordinal Utility Theory*. In circumstances of certainty, agents are conventionally taken to have rational preferences if their preferences are complete and transitive, and to choose rationally if their preferences are rational and there is no feasible option that they prefer to the chosen option. Many economists, in accordance with the theory of revealed preference, equate preference with choice definitionally (provided that consistency requirements are met). But identifying preference and choice cuts off too many questions. Keeping them distinct is useful both for clarity about what preference and choice mean, and to leave space for the formulation of alternative views of rationality that allow for the possibility of choosing an action that does not maximally satisfy preferences (see Sen 1973a, 1977a, and p. 688 below). This notion of rationality, as weak as it seems, may be too demanding, and there are a variety of weaker concepts, which we cannot discuss here.¹⁰

Completeness and transitivity together

¹⁰ See Sen (1971) and Edward McClennen (1990, ch. 2). Some of these views of rationality, as we shall repeatedly see—in the next footnote, for example—are more congenial to particular ethical theories than others.

establish a **weak ordering** of any finite set of alternatives, and one can assign numbers to options so that preferred options get higher numbers, and indifferent options get the same number. Any such assignment of numbers is an ordinal utility function. The *ordinal representation theorem* states that if an individual's preferences are complete, transitive, and continuous, then they may be represented by a continuous real-valued utility function (proved by Gerard Debreu 1959, pp. 54–59).¹¹

To be a utility maximizer is merely to choose the available option one most prefers (or, in the case of ties, to choose one among the set of maximally preferred options). Although the utility language was inherited from the utilitarians, some of whom thought of utility as a sensation with a certain intensity, duration, purity, or propinquity (Jeremy Bentham 1789, ch. 4), there is no such implication in contemporary theory. To speak of individuals as *aiming to* maximize utility or as *seeking* more utility may suggest misleadingly that utility is an object of choice, some ultimately good thing that people want in addition to healthy children or better television. But the theory of rational choice specifies no distinctive aims. This fact is of considerable importance to moral theory, for utility theory as such is detached from any hedonistic psychology.

It is worth emphasizing that, despite the fact that it recommends no substantive aims, utility theory remains a normative theory rather than a positive theory or merely a model or definition. It is not

¹¹ The assumption of continuity is arguably innocuous, but it is worth noting that at least one formal argument for utilitarianism (Marcus Fleming 1952) collapses without continuity of social or moral preference, and one formal argument for something akin to Rawls' difference principle turns crucially on rejecting continuity. See Steven Strasnick (1976, 1981); Peter Hammond (1976b); Claude d'Aspremont and Louis Gevers (1977).

a positive theory without further assumptions concerning the extent to which people are rational, and it is not merely a model or definition, because rationality is itself a normative notion. To define what rational preference and choice are, is ipso facto to say how one *ought* rationally to prefer and to choose. As utility theory places no constraints on what individuals may want, it has a much wider scope than economic theories.

2.1.2 *Expected Utility Theory.* Although the theory of rationality in circumstances of certainty is a central part of microeconomics, it is a thin theory. But the theory can be extended to cover cases of risk and uncertainty.¹² When actions do not lead with certainty to any particular outcome, they may be regarded as *lotteries* with outcomes as prizes. In offering a normative theory of decision making under risk and uncertainty, one then asserts that preferences among lotteries are complete, transitive, and continuous. In addition, one needs a reduction postulate identifying complex lotteries with simple lotteries that have the same probabilities and prizes (John Harsanyi calls it a “notational convention” 1977b, p. 24). Second (and last!) one needs the so-called “independence” or “sure-thing” principle, which says that if two lotteries differ only in one prize, A’s preferences between the lotteries match A’s preferences between the prizes. The (cardinal) representation theorem¹³ says that if an agent’s preferences are complete, transitive, and con-

tinuous and satisfy the reduction postulate and the independence principle, then those preferences may be represented by a utility function that possesses the expected utility property and is unique up to a positive affine (linear) transformation. A utility function possesses the expected utility property if and only if the (expected) utility of any lottery is equal to the utilities of its outcomes weighted by their probabilities. The representation theorem establishes that comparisons of utility *differences* are independent of the scale chosen. Although expected utilities are cardinal, they are not (at least without further argument) interpersonally comparable. As in the case of utility theory concerning choice under certainty, one relates choice to preference by asserting that individuals choose whatever feasible option is highest in their preference ranking—or, in the case of ties, a member of the maximal set of options. The probabilities in expected utility theory may be objective frequencies (as in von Neumann and Morgenstern 1947), but more commonly, they are subjective, representing the degrees of belief of individuals, and the axioms for rational choice under conditions of uncertainty imply that these degrees of belief will satisfy the axioms of the probability calculus. Moreover, it can be shown that if one’s degrees of belief do not satisfy the axioms of the probability calculus, then one can be led to accept a series of bets on some chance event leading to a certain loss whether the event occurs or not.¹⁴ Expected utility theory is thus not only a theory of rational preference and choice, but it may also provide a thin theory of rational belief.

2.1.3 *Questions about Utility Theory.* Expected utility theory imposes more structure on individual rationality than

¹² We follow Knight (1921), in speaking of situations of *risk* when outcomes have known probabilities and of situations of *uncertainty* when the probabilities of the outcomes of actions or even the range of outcomes is not known. See R. Duncan Luce and Howard Raiffa (1957, ch. 2).

¹³ See Harsanyi (1977b, ch. 3). proofs of similar theorems can be found in Israel Herstein and John Milnor (1953); Niels Jensen (1967); and John von Neumann and Oskar Morgenstern (1947). See also Frank Ramsey (1926) and Leonard Savage (1972).

¹⁴ See Frederic Schick (1986) for critical discussion of this so-called “Dutch-Book Argument.”

does utility theory under certainty, and it has proved much more controversial. Psychologists and economists have subjected expected utility theory to experimental tests, with results that are often unfavorable to the theory. Some results suggest not merely that people are not always rational, but that expected utility theory is flawed as a theory of rationality. These empirical anomalies, along with general critiques of expected utility (see Maurice Allais 1952; Daniel Ellsberg 1961; Sen 1985b; Levi 1986b; McClellenen 1990), have stimulated the formulation of several less demanding alternative theories of rationality under uncertainty (Farouk Gul 1991; Graham Loomes and Robert Sugden 1982; Mark Machina 1987, 1989; Amos Tversky and Daniel Kahneman 1979; Levi 1980).

Although we cannot survey the controversies concerning expected utility theory and its competitors here, they are directly relevant to moral philosophy as well as to economics. For expected utility theory is involved in arguments for moral conclusions, such as Fleming's (1952) Harsanyi's (1955), and William Vickrey's (1945, 1960) remarkable "proofs" of utilitarianism or David Gauthier's (1986) bargaining theory of justice. Alternative theories of rationality not only undercut these arguments, but in some instances, such as Levi (1980, 1986a), undercut the feasibility and attractiveness of specific moral views such as utilitarianism. Some of the more speculative of these alternative views of rationality, such as McClellenen's "resolute choice" (1990, esp. chs. 9, 12), offer new possibilities for conceptualizing and overcoming conflicts between individual rationality and social optimality. McClellenen defends the rationality of "resolutely" committing oneself to a course of action, such as cooperating in a series of prisoners' dilemma games (p. 719 below), and refusing (unless one is presented with unforeseen in-

formation) to make optimizing changes in one's plans. So resolute choosers might not defect in what they know to be the last round of an iterated prisoners' dilemma, and the "backwards induction" argument purportedly establishing the rationality of defecting on every move (p. 721 below) cannot get started. Faced with the possibility of dynamic inconsistency in one's actions (Robert Strotz 1956; Hammond 1976a, 1988), one need not adjust one's current decisions in the light of what one will prefer later. One can also decide on a course of action and stick to it.

Furthermore, although the issues concerning expected utility theory are highly theoretical, they are of immediate practical importance, as alternative theories may yield different recommendations about the rational thing to do. For example, in the case of nuclear power mentioned above (which obviously involves uncertainty, rather than merely risk), should we rely on our best guess as to the future consequences and maximize expected utility or should we recognize our ignorance of relevant probabilities and use some other principle of choice (Levi 1980)?

In addition to these controversies some moral philosophers have maintained that there are such things as irrational preferences, quite apart from any questions about consistency.¹⁵ Parfit, for example, holds that, "It is irrational to desire something that is in no respect worth desiring . . ." (1984, p. 123). He gives the example of a person who prefers a greater pain on a future Tuesday over a lesser pain on a Wednesday, merely because he or she does not care what happens on future Tuesdays. Parfit remarks, ". . . the fact that the agony will be on a Tuesday is

¹⁵ This is a classical view found in Aristotle and especially Plato. More recent defenders include Thomas Nagel (1970) and McDowell (1978).

no reason for preferring it. Preferring the worse of two pains for no reason, is irrational" (1984, p. 124). James Griffin (1986, chs. 3, 4) insists that certain classes of things such as accomplishment, pleasure, and "deep personal relations" should be among the objects of preference for all rational human beings, while other preferences are only rational if they can be shown to have some connection to such "objective" values. Such views are at odds with the way in which economists think about utility theory, though one might treat them merely as further constraints, rather than as a repudiation of the skeletal view of rationality espoused by economists.¹⁶

Indeed, even a philosopher such as Gauthier, who builds a theory of value on expected utility theory and who is a forceful critic of an objective view of value, insists that expected utility theory is too weak. Gauthier argues (1986, ch. 2) that rational preferences should reflect experience and survive reflection, and that the preferences one reveals in behavior should match the attitudes one enunciates verbally. On the basis of criteria such as these, Gauthier distinguishes values from mere tastes.

Accepting utility theory also gives rise to a deep problem for both prudence and morality whenever preferences change. As Richard Brandt (1979, ch. 13) and Parfit (1984) point out, to consider the good for an individual requires more than attention to his/her given preferences, no matter how purified. For if a given preference ranking changes, then it is unclear whether an individual is made better off by satisfying or frustrating the original preferences.¹⁷ It seems that an accep-

table view of rationality and welfare will have to look beyond, behind, or beneath given preferences (Stewart 1989). This point has clear policy relevance if institutions or policies have systematic effects on the preferences people form. For then assessment of policy requires either deciding which preferences to promote or deciding on procedural grounds that certain kinds of institutions provide a suitable framework within which desirable preferences will develop (McPherson 1982, 1983b). One should also consider whether there are conflicts between first-order preferences (for example for cigarettes) and so-called "metapreferences" or preferences for preferences (see page 688 below).

2.2 Incorporating Moral Behavior into Economic Theory

We can now discuss how moral features of conduct might be related to the standard view of rationality. A useful starting point is with the recent literature that attempts to incorporate moral dimensions of behavior into positive accounts of economic phenomena.¹⁸ This

¹⁸ Some economists have adopted the view that purported explanations of economic phenomena that presuppose non-self-interested conduct are necessarily superficial and ad hoc. An adequate explanation must show that the apparently moral behavior can, on deeper analysis, be shown to be consistent with—and indeed can be derived from—self-interest. This position is examined by Jon Elster (1979, ch. 3). The research strategy of which this is an example is, of course, much more general in economics. Many contributions to macroeconomics, for example, have been similarly criticized as ad hoc unless they can be derived from the familiar equilibrium model of economics (Douglas Wade Hands 1988). About this general strategy it is perhaps worth making these three observations. First, at any stage in the development of economic understanding, it seems rash to reject theories outright because they appear superficial or ad hoc, simply on the hope that more adequate theories will become available later. Second, there cannot be any general guarantee that the attempt to reduce superficial theories to ones that are consistent with familiar microeconomic assumptions will work—we have no a priori guarantee that this set of axioms is true and complete. But third, it is per-

¹⁶ The relation between preference satisfaction and well-being is further discussed in Section 3.1.1 below.

¹⁷ Indeed, one faces within an agent's life something akin to the problem of interpersonal utility comparisons (Gibbard 1986).

literature draws eclectically on generalizations about moral behavior from sociology, history, observation, or even experiment to develop modified theories about the behavior of households, firms, and hence economies. Many of these studies have emphasized that moral constraints on behavior may be *instrumentally* important to good economic performance, although one can also study how moral practices might hinder economic development (Diego Gambetta 1988; Edward Banfield 1958). And the effects of moral commitments on the character of economic outcomes can be of interest even if there is no effect on efficiency.

2.2.1 *The Effects of Moral Commitments.* Moral norms may have significant impacts on the provision of public goods, voting, and activities involving the production of externalities (Michael Taylor 1987; Thomas Schelling 1978; Jane Mansbridge 1990; and Elster 1989a). But perhaps the most sustained and illuminating treatments of moral influences on economic behavior have been concerned with work motivation and the organization of the firm. Since at least Marx and Mill, economists have recognized that worker performance is closely related to workers' perception of the legitimacy of authority and the fairness of distribution.

Recent research on these issues begins with the recognition that contracts for labor are unavoidably incomplete, and monitoring of worker performance is imperfect and costly. Any sort of performance-based contract is subject to moral hazard problems on both sides—with workers attempting to supply less effort than is agreed and employers trying to

understate worker performance levels. A variety of devices for overcoming these problems have been examined, some of them relying strictly on self-interested motivation by all parties (Akerlof and Janet Yellen 1986). But both everyday experience and a substantial body of sociological research suggest that norms and moral dispositions have an important influence.¹⁹ Akerlof (1983, 1984) has emphasized the role of norms of fairness—"a fair day's work for a fair day's pay"—in governing effort levels, and argues that sustaining these norms requires that workers and employers view part of the exchange between them as an exchange of "gifts"—a gift of work effort above what can be enforced and a gift of pay above what the market demands (see also Harvey Leibenstein 1976, 1979). A related but distinct point is the role of *trust* in governing worker-employer relationships (Roland McKean 1975; Arrow 1974; Gambetta 1988). Trust is particularly important in long-term work relationships, where the rewards for reliable service may be long delayed. Commentators on the Japanese economy have identified distinctive cultural norms regarding cooperation and work effort as important to the functioning of Japanese firms (Ronald Dore 1987; Michio Morishima 1982).

Economists have also noted that moral norms influence the comparative efficiency of alternative ways of organizing the productive enterprise (Oliver Williamson 1980; Alan Fox 1974; Hans Nutting 1976; McPherson 1984). Workers' rights to control over work decisions and workers' beliefs about distributive fairness have been alleged to play an important role in determining the relative

fectly reasonable for some economists to devote energy to searching for such reductions in particular cases: sometimes such reductions may work, and, if persistent efforts fail, the efforts may constitute grounds to take the apparently ad hoc claims seriously (Hausman 1992, esp. ch. 6).

¹⁹ George Baker, Michael Jensen, and Kevin Murphy (1988) provide a helpful review of observed labor market phenomena that are hard to explain on standard economic assumptions and that seem to indicate an important role for moral norms in influencing the labor market.

productivity of entrepreneurial and worker-controlled firms. The general proposition, which dates back to Mill, is that workers will be more willing to supply effort in circumstances where they have a say in the rules governing the enterprise and where there is good reason to believe that a strong common effort will result in larger rewards for members of the group (Louis Putterman 1984; Samuel Bowles 1985; McPherson 1983a; Michael Reich and James Devine 1981; Bowles and Herbert Gintis 1990).

Sociological and normative aspects of work relations may also influence labor markets.²⁰ Indeed, social norms in the workplace may help explain “sticky wages” and persistent unemployment. “Gift exchange” or more generally normatively governed wage relations explain the existence of a wage that persists above the market clearing equilibrium as a means of providing workers with an incentive to perform well on the job (Akerlof 1984; Solow 1981; Alan Blinder and Don Choi 1990; Bowles 1985). Such an “efficiency wage” may explain a persistent excess supply of labor, but the persistence of such a wage itself raises problems. Why don’t unemployed workers seek to undercut this wage by offering to work for less than those with jobs? Why don’t employers actively seek such workers? Solow (1990, ch. 2) and Jørgen Weibull (1987) suggest that further social norms come into play—a norm among workers against bidding against their fellows for jobs and a norm among employers against seeking to undercut their existing workforce.

These lines of analysis point to the importance of normatively governed restraints on individual maximizing in explaining the performance of firms and the

behavior of labor markets. They also raise further questions about how such norms originate and persist and about whether adherence to the norms is economically efficient and/or morally justified (Elster 1989a, 1989b; Sugden 1989).

If moral dispositions of economic agents influence outcomes, then economists concerned with public policy and social welfare may want to say something about whether public policies ought to encourage certain kinds of moral behavior. Attending empirically to moral dimensions of individual behavior may thus affect economists’ evaluation of institutions or policies. A simple example is provided by Melvin Reder’s (1979) analysis of the economics of honesty. Reder proposes to measure a person’s honesty by the size of the reward required to induce him or her to lie.²¹ *Ceteris paribus*, more honesty will reduce contracting and policing costs, and therefore policies that promote honesty will raise net economic output. Reder suggests that social investment in schools and churches may have this economic justification.²²

This analysis makes the value of moral dispositions and of institutions which promote them instrumental to given eco-

²¹ There is an anecdote, whose origins we have been unable to trace, of someone who attempted to bribe Abraham Lincoln. Lincoln kept brushing him off genially and the briber kept increasing his price. When the price got very high, Lincoln clouted him. When asked why he suddenly got so aggressive, Lincoln responded—because you were getting close to my price! It would be hard to model this sort of honesty in Reder’s way.

²² The instrumental or functional *justification* of some moral dispositions, or of institutions that promote morality, does not, without much further argument, *explain* their existence. As Elster has argued, pointing to the favorable consequences of a social practice does not explain its existence or persistence in the absence of an account of the mechanisms by which the favorable consequences produce or sustain the practice. In biology, natural selection provides such a mechanism. In the social sciences functional explanations are more dubious. (See G. A. Cohen 1982; Friedrich von Hayek 1988; Elster 1989a; Viktor Vanberg 1986; and Geoffrey Hodgson 1991.)

²⁰ As Robert Solow (1990) emphasizes, this thought has a distinguished history in such works as Alfred Marshall (1887) and John Dunlop (1957).

conomic ends. The advantages of embedding moral behavior into the familiar scheme of economic evaluation in this way are substantial, and this approach is compatible with influential philosophical accounts of the role of morality in social life. Philosophers within both utilitarian and contractualist traditions have argued that moral restraints are justified by the fact that the pursuit of self-interest may lead to nonoptimal social outcomes (Kurt Baier 1958; John Mackie 1977; Gauthier 1986).²³

There are various ways in which moral beliefs and dispositions might be assessed instrumentally. Perhaps the simplest way would be to take economic efficiency to be the relevant goal (as Reder does). Yet the moral dispositions of economic actors may also influence distributive outcomes, so equity as well as efficiency judgments may figure in instrumental assessment of moral conduct. Questions of power also intrude, as would be stressed by radical economists who have argued that the tendency for workers to "shirk" under capitalist productive arrangements is a justifiable reaction to the oppressive circumstances of their work (Harry Braverman 1974; Richard Edwards 1979).

Viewing moral rules as promoting social welfare is useful in linking them to familiar modes of economic analysis, and it has also given rise to striking evolutionary speculations (Gibbard 1990). Yet this way of looking at moral rules broaches deep questions about the relation between rationality and morality. As McKean (1975) notes, this instrumental outlook takes moral dispositions as means to advance the satisfaction of given preferences. But where does the inclination

²³ Historically, David Hume's analysis of justice is an important progenitor of this tradition. And, indeed, it can be argued that Hume's young friend Adam Smith gave substantial credit to this view of morality both in *The Theory of Moral Sentiments* and *The Wealth of Nations*.

to follow moral rules fit into the actors' preferences? If it is part of their preference system, then there seems to be an element of circularity in influencing the strength of the inclination (preference) for following moral rules as part of a strategy for maximizing overall preference satisfaction. But if moral dispositions are not part of the preference system, then the fact that a general practice of following the norms serves a useful social function leaves open whether it is *individually* rational for persons to adhere to moral norms.

2.2.2 *Modeling Moral Commitment.* These conundrums lead back to conceptual issues concerning the relations between morality and rationality. The central question is whether there is any inconsistency between economically rational action and morally motivated conduct. There seems offhand to be no inherent incompatibility between utility theory and the demands of morality. For utility theory places no constraints on the objects of preferences. Preferences can just as well reflect the dictates of conscience as the promptings of the libido.²⁴

Despite the flexibility of utility theory, it is questionable whether the full range of familiar moral phenomena can be accommodated within it. Transitivity and completeness require that people have a clear view of what is the right thing to do in every situation and of how much weight to give moral concerns in every choice. As problems of moral backsliding, of weakness of will, and of doubt are familiar features of everyone's moral

²⁴ It is, however, important to appreciate that the objects over which preferences are defined may have to be quite different from the familiar range of economic commodities if the preferences are to rationalize moral choices. To cite a trivial example, a delicious piece of apple pie that was known to be stolen would not be the same object as a physically similar piece of pie that came as a gift. Susan Hurley (1981, part I) argues that values inevitably constrain which descriptions of the objects of preference are "eligible" (and that consequently values and preferences are interdependent).

experience, it is hard to believe that the choices of actors who are influenced by moral concerns can be rationalized by a complete and transitive preference function, nor is it obvious that it is reasonable to demand that they can be.²⁵

But even if utility maximization and morally motivated conduct are formally consistent, one might still question how helpful utility theory is in analyzing moral conduct. As Sen (1977a) has argued, even though utility theory may be too demanding in some regards, it may at the same time impose *too little* structure on problems of choice to allow productive analysis of moral conduct. For example, the question with which we concluded the last section—Is moral norm-following consistent with individual rationality?—is drained of significance if people's moral concerns are automatically built in to their preference functions. The vocabulary in which preferences and choices are described needs enrichment.

One obvious step in this direction is to distinguish *self-regarding* or *egoistic* from *other-regarding* or *altruistic* preferences. No special formal difficulties arise in including either the utility or consumption levels of other individuals in a person's utility function.²⁶ Terminologically, it is useful to mark the distinction

²⁵ There is an interesting current literature questioning whether a specifically deontological moral system (one that specifies what is right or wrong without regard to the consequences) can be consistent with utility theory (John Broome 1991b). Many question whether a truly deontological ethics—particularly one employing “agent-centered constraints” (p. 705 below) is defensible (Samuel Scheffler 1982, 1988; Shelly Kagan 1989).

²⁶ Arrow (1972), Gary Becker (1981), Kenneth Boulding (1973), David Collard (1978). There is no reason in principle why the sign on other people's utility has to be positive. Things are particularly simple if what enters one person's utility function is the level of satisfaction of the *self-regarding* preferences of someone else. If instead what enters is the other person's overall level of satisfaction, including satisfactions obtained from the welfare of others, tricky strategic issues arise owing to the feedback effects between different people's welfare (see Elster 1989a).

by referring to the satisfaction of self-regarding preferences by some term such as “personal interests” or “standard of living,” recognizing that this differs from the person's overall level of satisfaction, including the satisfactions of sympathy.²⁷ However, not all nonegoistic concerns are well-captured by such notions of sympathy or altruism. First, one may be concerned not only with the level of another person's welfare or consumption, but with how it is attained. One may wish to be an active contributor to another's well-being. As Arrow has noted, the desire to contribute may have substantially different implications than mere altruism. If many affluent people prefer a higher level of consumption for “the poor,” then the poor's consumption level is a public good to the affluent. If instead each wishes to contribute to helping the poor, the public-goods problem goes away (Hochman and Rodgers 1969). More generally, this *agency* aspect of individual motivation—the desire to *do* certain things rather than simply to enjoy the consequences of their being done—has important implications in modeling preference and choice (Sen 1982b, 1987a).

Sen proposes another distinction between sympathy and “commitment.” Consider the example of a soldier on sentry duty. Her motives in staying alert may include a self-interested fear of punishment and a sympathetic concern for the welfare of those she is guarding. But she may be motivated as well by a desire to do her duty, to live up to a code she has acknowledged as binding on her. Sen proposes to define “commitment” in terms of her willingness to undertake an

²⁷ There is no canonical usage here. For a helpful discussion of alternative ways of distinguishing among the various categories of matters “an agent has reason to favour or promote” (Williams 1987, p. 95) see Sen's essay, Williams's comment, and Sen's reply in Sen et al. (1987). What is important is not the precise names given to different concepts here, but an awareness of distinctions among those concepts.

action (in this case staying alert) even if her welfare—including both self-interested and sympathetic components—were reduced by doing so.²⁸ Economists schooled by the theory of revealed preference may be tempted to insist that if the agent chooses to stay alert, then she must have preferred to; and staying alert makes her better off. But identifying preference, choice, and welfare begs too many important questions. Even though it may be possible formally to model the committed agent as maximizing utility, it seems enlightening not to do so. For then one can begin to consider just what sort of role morality might play in constraining choices, and one can recognize the special problems of self-control and weakness of will to which duties give rise.

Should a person who acts against her narrow personal interests out of sympathy or duty be said to act irrationally? Elaborating the structure of internal motivations has the helpful effect of making that question dissolve. If one stipulates that acting rationally requires acting in accord with selfish preferences, then the answer is “yes.” If sympathetic but not “committed” concerns are captured in what the agent maximizes, then the answer concerning commitment is still “yes.” But if one says most generally that acting rationally means acting on reasons that seem to be good ones, then the answer is “no.”²⁹ It is important to be able to draw these distinctions, if only because the alleged “irrationality” of moral commitment needs to be distinguished from

the irrationality of passion, vendetta, or schizophrenia.

The preceding pertains to making sense of the actions of a person with a set of complex but presumably settled motivations. But one may also need to make sense of the phenomena of reflecting on, assessing, and sometimes revising one’s motives or reasons for acting. The devices of multiple preference systems and of higher-order preferences may be particularly useful in this regard. Multiple preference systems very naturally model internal conflict concerning such personal choices as whether or not to smoke (Schelling 1984, pp. 57–112) as well as more clearly moral choices such as whether to contribute to charity (Ami-tai Etzioni 1986, pp. 177–80; Harry Frankfurt 1971). Indeed some scholars have argued that the choices of morally influenced persons are in general better modeled through a multiple-utility framework (Howard Margolis 1982; Etzioni 1988). However, this approach also raises difficult questions, for it is not clear within the multiple-utility framework how to explain which ranking scheme will prevail in which circumstances. Indeed, if that determination is itself made by a consistent preference scheme that exists at a higher level, then it may make sense to reinvoke standard utility theory at this higher level.

The “homo economicus” of contemporary economics is “homo rationalis.” This fact embodies a commitment both to a certain mode of modeling and to a normative theory of prudence. The theory of rationality is already a fragment of a theory of morality. But the view of rationality economists endorse—utility theory—may not even be compatible with moral behavior, and it does not provide a rich enough picture of individual choice to permit one to discuss the character, causes, and consequences of moral behavior. Of course economists need not

²⁸ Notice that this definition (like Reder’s view of honesty, p. 685 above) involves a subjunctive. Duty or more generally committed behavior may happen to coincide with personal welfare; but the defining characteristic is that the committed action would be carried out *even if* this were not so.

²⁹ Or anyway, “not necessarily.” Clearly a person could be misinformed or confused and be irrational for that reason at the same time she was being altruistic or dutiful.

aspire to provide a general theory of human action. Yet they should not shrug their shoulders at the difficulties in meshing moral behavior with economic rationality. For it seems that moral behavior may have important consequences for economic outcomes, and the propagation of utility theory has moral implications.

3. Evaluating Economic Arrangements

Distinguishing better from worse economic policies and outcomes is a central purpose of normative economics, and one that inescapably relies on moral judgments. Moral evaluations of economic arrangements are built into welfare economics and into the terms of policy debate. Because they permeate economics, it is worthwhile to make these moral appraisals as sensible as possible.

3.1 Dimensions of Evaluation

When people think of moral philosophy, they think of attempts to lay out and apply a general framework of moral principles and methods. This is a reasonable picture of landmark work such as Kant's ethics or utilitarianism, and of many modern theories, including Rawls' theory of justice, Nozick's entitlement theory, and neoclassical welfare economics. Because such theories are broad and systematic, and because each theory provides its own interpretation of important concepts such as "well-being" or "equality" or "rights," such theories are difficult to analyze and compare. In our view, the parts of ethical theories we shall survey here may be of more interest to economists than are the whole systems.

3.1.1 Utility Theory and Human Good. All plausible moral views assign an important place to conceptions of individual good, welfare, or well-being. This is most obviously true of utilitarianism, which takes morality as maximizing some

function of the welfare of individual members of society. But even nonutilitarian views that emphasize notions of rights, fairness, and justice need a conception of human well-being. These views not only recognize the virtue of benevolence, which requires some notion of human good, but even their core notions make reference to well-being. Thus, for example, justice or fairness is understood in terms of treating the interests of different persons properly, and acting rightly will often involve avoiding harm to other individuals. And notions of harm and of interest are plainly connected to notions of human well-being.

Given economists' commitments to utility theory in explaining human choices, it is natural that they would look to levels of utility or of preference-satisfaction as the fundamental measure of human well-being for evaluative purposes as well.³⁰ In applied work, it is true, economists often rely on more objective measures of "real income" rather than utility measures. This, however, is typically viewed simply as a compromise with data limitations: real income is regarded as a proxy for preference satisfaction (and a highly imperfect one at that—see John Hicks 1940; Samuelson 1950a; Sen 1976b; and Jan de Van Graaff 1957).

However, it is questionable whether preference satisfaction is an adequate conception of individual well-being. Measurements of well-being have many purposes—to establish benchmarks for equality, to assess progress over time, to compare living standards across communities, to weigh the comparative

³⁰ If these explanations of choices are, as they appear to be, accounts of the reasons of rational agents, then what is chosen must, given the agent's beliefs, serve his or her interests best. If one assumes, as is common, self-interest and full information, the theoretical temptation to identify well-being with the satisfaction of preferences becomes overwhelming. Concerns about rationality link positive and normative economics (Hausman 1992, chs. 4, 15).

claims that different persons may justifiably make on social resources, and so on. Different conceptions may suit different purposes, and for some purposes central to moral reasoning, preferences may not form a suitable basis (Sen 1987b, 1987c).

A variety of reasons have been advanced for why, quite apart from measurement difficulties, preference satisfaction cannot serve some of these purposes well. First, there are reasons to doubt whether there is any context at all in which one can correctly identify well-being with the satisfaction of preferences. For preferences may be based on false beliefs. One does not make a person better off by satisfying a preference for a glass of poison based on the false beliefs that it is water or that it will make one immortal. Most philosophers who associate well-being with preference satisfaction consequently require that the preferences be suitably "corrected." This introduces an objective element of well-grounded or true belief into a preference-based standard.³¹

But problems remain. For preferences that are not based on indisputably false beliefs may nevertheless be idiosyncratic or based on highly contestable beliefs. Even if the preferences of a starving person who begs for money to make a sacrifice to his god are well informed, one may not believe that satisfying them will make this person better off. If the notion of individual well-being is to be employed in weighing the strength of claims for social provision, it must permit reasoned social agreement on basic components of well-being and on the relative

"urgency" of claims to different goods (Thomas Scanlon 1975). One might thus acknowledge a moral obligation to feed the person, but not to finance his worship.³² Notice that a subjectivist might claim that this resistance to honoring idiosyncratic preferences is purely practical: we would open ourselves to all kinds of manipulation and misrepresentation if we let people's subjective reports of their needs govern public distribution of benefits.³³

There are further reasons to question whether the conception of well-being as the satisfaction of preferences is suitable for assessing claims to scarce resources. For example, those with expensive tastes may make unfairly large demands. A person who has cultivated a taste for "pre-phylloxera claret and plover's eggs" (Arrow 1973, p. 254) may be miserable without them and, in one sense, worse off than someone in similar circumstances who delights in beans and franks (see also Ronald Dworkin 1981). But one may question whether social policy should be responsive to such preferences. Defenders of a preference-satisfaction view may argue that what is at issue is fairness rather than the conception of well-being: the person is worse off without the plover's eggs, but policy makers don't have to care. In reply, defenders of more objective views of well-being can simply deny that the fancy eater's standard of living or level of well-being is lower at all; there is no need to explain away our indifference to his or her "plight." Similarly, as Harsanyi among others has argued, it is questionable whether satisfying racist, sadistic, and other such

³¹ The satisfaction of some preferences, such as preferences for the prosperity of comparative strangers, seem to have nothing to do with an individual's own well-being and many have argued that other-regarding preferences should not be included in our understanding of individual well-being. (See Brian Barry 1965, pp. 11–15, 61–66, 71–72; Ronald Dworkin 1977, pp. 234f; D. W. Haslett 1990; and Thomas Schwartz 1982.)

³² Scanlon (1975). See also Mark Sagoff (1986). One can see here one reason for preferring in-kind provision to transfer payments.

³³ One might also object that it is paternalistic not to respect the worshipper's preferences. But paternalism and antipaternalism are views about what it is permissible to do; they are not views about what individual welfare is. See footnote 8 above.

antisocial preferences should count as contributing to individual well-being.³⁴

A final difficulty concerns the effects of culture on preferences and the problem of “adaptive preferences” (Elster 1983; Sen 1987b, 1990). For example women who are systematically denied roles in public life or equal shares of consumption goods may learn (à la Sinead O’Connor) not to want what they have not got. Their utility may be reasonably high, but objectively their standard of living or level of well-being may be low.

The difficulties with a subjective preference-satisfaction view of well-being have elicited two broad types of responses. One response has been to hang on to preference-satisfaction, and, on the one hand, to deflect or defeat some of these objections or, on the other hand, to “launder” preferences of these disagreeable features. Whether these responses succeed is a controversial question.³⁵ An alternative strategy is to propose a different conception of well-being. Not only are there traditional hedonistic views and hybrids such as Sidgwick’s (1901, bk. 2, ch. 2), but there are also “perfectionist” views (Griffin 1986, ch. 4; Parfit 1984, ch. 6; Joseph Raz 1986, ch. 12) in which the desires that count and their weight depend on the objective

³⁴ 1977a, p. 56. One may be able to deflect the claims of those with expensive or antisocial preferences without questioning the identification of well-being and the satisfaction of preferences. Not only can one claim (as noted in the text) that some moral issue distinct from the concept of well-being is involved, but one can also note that preferences are malleable, and expressions of preferences respond to incentives. The frustration of expensive and antisocial preferences may contribute to changing them and would discourage their expression, which would enable more preferences to be satisfied.

³⁵ Griffin (1986) provides an extended defense of an informed preference satisfaction view of individual well-being. For a critical summary of arguments against the preference-based approach to measuring well-being, see the introduction to Sen and Williams (1982), and for a powerful general reply, see Arneson (1990).

value of their objects. Indeed what Parfit calls “objective-list” views (1984, Appendix I) jettison most links to agent’s desires or aims.

For an objective view that may be more relevant to economists, consider John Rawls’ *A Theory of Justice*. Although Rawls sees well-being as the satisfaction of rational preference or desire, he does not regard this conception as appropriate for the purposes of a theory of justice. Rawls proposed instead that well-being be measured by an index of “primary social goods.” Goods such as education or income are all-purpose means or, as Rawls put it, things that you want, whatever else you want (1971, p. 92). Rawls does not regard primary social goods as proxies for utility levels. On the contrary, in his view they offer an alternative basis for a more settled social agreement on what is important to well-being (1982). Rawls’ approach avoids the expensive tastes and antisocial preferences problems and, he argues, also provides a more impartial perspective for comparing how well-off different individuals are than a preference standard does.

Amartya Sen (1985c, 1987b, 1987c, 1992) criticizes the primary-goods approach because it focuses on the external means that permit people to attain various functionings rather than on “capabilities,” or abilities to achieve functionings.³⁶ In focusing on external goods it leaves out more internal features of persons’ circumstances, such as physical handicaps, which may profoundly affect well-being. In focusing on *primary* social goods, Rawls’ account also may not work well for comparing the well-being of persons who are much above the mini-

³⁶ A capability is the ability to achieve a certain sort of functioning—literacy is a capability; reading is a “functioning.” People may value capabilities for their own sake as well as for the functionings they permit—you’re glad to know you can walk around even if you are inclined to stay put.

mum. Sen's alternative is to define well-being in terms of the set of "functionings" a person achieves. Being well nourished is an example of a functioning whose achievement may be impaired either by internal obstacles (such as a digestive disorder) or external ones (a lack of money to buy food). Functionings are of course a vector (being well nourished, reading books, sleeping well) and combining them into a measure of well-being raises a serious index-number problem, the solution to which requires settling on the relative weight to be given to different capabilities and functionings.³⁷

Even if the problems of weighting and measuring the components of well-being are greater for these more "objective" approaches than for standard utility theory (which is not clear), it can still be argued, following Sen, that "it is better to be vaguely right than precisely wrong." In particular, these approaches focus research on well-being in different directions, which link up more naturally to the normative terms in which policy is debated.

A further important advantage of more objective approaches is that they avoid the problems of interpersonal comparison that derive from identifying well-being and utility (construed as an index of preference satisfaction). Formerly, economists such as Pigou argued that overall welfare would be maximized by equalizing incomes as much as was consistent with retaining incentives to produce. Citing diminishing marginal utility of in-

come, they maintained (not implausibly), that an extra one hundred dollars contributes less to the well-being of someone with an income of \$50,000 than to the well-being of someone with an income of \$5,000. Other things being equal, a more equal distribution of income would increase total welfare.

Whatever merits this argument may have if welfare is interpreted as some sort of general material well-being, it has (without further ethical premises) little worth if welfare is interpreted as preference satisfaction (Robert Cooter and Peter Rappoport 1984). Lionel Robbins argued compellingly that because preference rankings are entirely subjective and introspective, there is no way to compare where *A* is in *A*'s preference ranking with where *B* is in *B*'s ranking (1935, ch. 6). The fact that people nevertheless make interpersonal comparisons of well-being depends, Robbins argues, on the value judgment that the preferences of people in similar circumstances are similarly satisfied. (One might—in our view more plausibly—instead question whether these comparisons of well-being are comparisons of preference satisfaction.)

Yet Robbins did not convince everyone. Adopting a view similar to Arrow (1978) and Serge-Christophe Kolm (1972), Harsanyi (1955 and especially 1977b, ch. 4), for example, has argued that there is, in effect, a single utility function that depends not only on the objects of choice, but on causal variables affecting and in effect identifying the chooser. Because psychological theory cannot yet specify this underlying utility function, one employs one's empathic abilities to determine one's "extended" preferences between ordered option-agent pairs such as $\langle x, \text{Norma} \rangle$ and $\langle y, \text{Grover} \rangle$. Judgments resulting from putting yourself in someone else's shoes in this way can be used to construct inter-

³⁷ This issue of weighting might seem to bring preferences in through the back door—ultimately, it might be said, the importance of different functionings would depend simply on the weight they are given in people's subjective preferences. However, while people's preferences for functionings and the capabilities they presuppose are relevant to their weight, it is by no means clear that a reasoned public discussion over these weights would wind up concluding that individuals' subjective preferences about them were their only or most important determinant.

personal comparisons.³⁸ It is also possible to find a foothold for making interpersonal comparisons of utility if one interprets utility in psychological terms, rather than as bare preference. Thus, if a high utility level corresponds to a certain state of mind (or set of possible states of mind), it may be possible to make sense of at least partial interpersonal comparability (Sen 1979d; see also Elster and Roemer 1991).

The identification of well-being with the satisfaction of preferences is questionable in itself. It gives rise to intractable problems of interpersonal comparisons, it is biased by adaptive preferences, and it fails to link up with the normative terms of policy debate. Moreover, the standard notion of utility is not easily measurable, and even if it were, economists cannot defend a measure of individual well-being as a good operational proxy without having some conception of what it is supposed to be a proxy for. The only real advantage of identifying well-being with the satisfaction of preferences is that it apparently links welfare problems immediately to fundamental economic theory. But the link is bogus unless the theoretical "welfare" problems are truly problems of human welfare or well-being.

3.1.2 Liberty and Rights. The concept of human well-being is crucial in moral theory, but it is not the only morally significant concept; and even if economists had a more cogent construal of well-being, they would still be hard pressed to evaluate economic outcomes without bringing in other notions.

One of these other important notions is liberty. Many people, including many economists, have the conviction that the protection of certain liberties is of great moral importance regardless of whether

such protection always increases well-being. For example, many object to paternalistic laws, such as laws requiring the use of seat belts, as infringements on freedom regardless of whether these laws increase welfare. Social deliberation, unlike welfare economics, often treats protecting freedom and enhancing welfare as independent goals.

It is ironic that welfare economics focuses almost exclusively on the Pareto efficiency concepts. For economists typically value individual freedom; and much of the traditional case for capitalism was not so much in terms of its capacity to "deliver the goods" as in terms of the protection that the separation of economic and political power offers to individual liberty (Hayek 1960; Friedman 1962). Economists favor effluent taxes over direct state regulation of pollution not only because the taxes are purportedly Pareto superior to the regulation, but also because the regulation is a threat to freedom. Yet the efficiency case has been a part of "scientific" welfare economics, because it has appeared to rely only on uncontroversial moral premises, while the argument in terms of freedom has been for essayists or ideologues, because its moral premises have appeared more "philosophical."

To the extent that concerns about liberty underpin economists' normative judgments, they should be incorporated systematically into economists' methods for evaluating states of affairs. Doing so raises problems of *definition*, of *moral justification* and of *integration*. In other words, (1) What is freedom? (2) How can claims about the moral importance of freedom be defended? and (3) How can liberties be incorporated along with other values into a coherent scheme of ethical evaluation?

The term, "freedom," is notoriously slippery. Isaiah Berlin (1969) distinguished what he called "negative" from

³⁸ For a critique of this "mental shoehorn" tactic, see Alfred MacKay (1986) and Griffin (1986, ch. 7).

“positive” liberty. Negative liberty is liberty from interference, the freedom to act without deliberate negative social sanctions. Berlin defined positive liberty as autonomy or self-determination, while others have thought of positive liberty more in terms of the range and quality of alternatives open to an individual.³⁹

It is hard to separate a concern with negative liberty from concerns with freedom in a wider sense. Rawls has attempted to cope with this complexity by distinguishing “liberty,” understood as a formal notion, from the “worth of liberty”; others have distinguished “formal” from “effective” liberty. Sen’s capabilities (p. 691 above) provide one construal of a sort of positive freedom. Autonomy, which is often seen as the central notion of positive liberty, is a complex notion that draws on other notions of freedom.⁴⁰ Moreover, one can distinguish between autonomy as an achievement and the capacity for autonomy (Raz 1986, pp. 204f). Two concepts of liberty are not enough to capture all the dimensions of interest.

Why is liberty in these several senses morally important? Negative liberties are important because they help people get what they want. They promote not only welfare but also integrity, dignity, and autonomy itself. One can make a similar instrumental case for positive freedoms. Some classic defenses of liberty, including Mill’s (1859) and Hayek’s (1960), are almost entirely instrumental: coercion is costly and dangerous and likely to be misapplied. In the long run, progress de-

pends on individuality, which can only flourish in circumstances of liberty.

Although one can thus value freedom, without regarding it as intrinsically valuable, many would go further and maintain that positive freedom, construed as self-determination or autonomy, is itself intrinsically valuable (Griffin 1986, ch. 4; Raz 1986, ch. 15; Sen 1985a, 1985c). Indeed, autonomy, as rational self-legislation, is, in Immanuel Kant’s (1785) view, perhaps the central moral value. The fact that liberty, and, relatedly, the prevention of rights violations, may be valued intrinsically as well as instrumentally, leads to complications in including liberty and rights in overall schemes of ethical valuation. We will return to these issues after saying something about the definition and justification of rights.

Although rights and justice obviously contribute to welfare, many people, including many economists, believe that the moral force of rights is not exhausted by their instrumental role. Even if permitting slavery might somehow make us richer, such a practice violates rights and is not to be tolerated.

Rights are often linked to liberties—and for good reason. For one kind of right (such as the right of ambulances to speed) simply is a liberty, and liberties constitute an important part of standard “claim-rights.”⁴¹ *P*’s (claim-)right to *X* involves both liberties of seeking, acquiring, possessing, or using *X* along with a set of duties on others, which may include positive duties such as performing a contracted service as well as negative duties of noninterference. Legal rights define what a person is permitted to do without incurring a threat of legal punishment, what personal claims are legally protected, and consequently what other per-

³⁹ In speaking of the quality of alternatives, one brings individuals’ welfare back into the picture. George Stigler argued on this basis that “liberty” and “wealth” are really synonymous—both index the quality of your options assessed in welfarist terms (1978). But there is a real loss in liberty if you are locked in your house for the day, even if there is no place you would rather be (see Elster 1983 and Sen 1987a, ch. 3).

⁴⁰ Feinberg (1986, ch. 18) distinguishes many different aspects of autonomy and draws a helpful analogy to national sovereignty.

⁴¹ The terminology is from Wesley Hohfeld (1923). For general discussions of rights see Feinberg (1973, chs. 4–6) and the introduction to Jeremy Waldron (1984).

sons or governments are prohibited from doing or are required to do. Moral rights similarly define permissions, guarantees, prohibitions, and requirements as matters of moral principle rather than law.

Liberties are important not only to the definition of rights, but also, at least within the Kantian tradition, to the specification and justification of rights. In the view of philosophers such as Herbert L. A. Hart (1955), the limitations imposed on others by *P*'s claim-rights must be justified by their contribution to *P*'s freedom. Rights determine the distribution of freedom. The libertarian considerations that justify placing control with *P*, justify the duties the right implies for others. Even if trespassers on *P*'s property pay fees that restore *P* to the same level of welfare, they have on this view still violated *P*'s property rights. "Isn't the very idea of a forced transfer contrary to the autonomy or liberty thought constitutive of rights?"⁴²

But many philosophers would assert that other interests besides *P*'s interests in liberty can also ground *P*'s rights and the consequent duties of others. There is an intricate consequentialist case to be made for moral rights going back to Sidgwick and Mill. (For relevant contemporary work, see Roy Harrod 1936; Rawls 1955; David Lyons 1977; Donald Regan 1980; and Russell Hardin 1988.) From a utilitarian perspective, rights serve as rules that should be crafted to promote welfare. But even nonutilitarians often ground rights in terms of a wider set of

individual interests. For example, in Raz's view, "if an individual has a right then a certain aspect of his well-being is a reason for holding others to be under a duty" (1986, pp. 171–72). Such a "benefit" view of rights is compatible with a utilitarian account, but it leaves open the question of when an interest of a potential rights-holder is sufficient to justify imposing duties on others. Rights may, depending on the circumstances, justify imposing a variety of duties on different individuals. Consider, for example, the array of duties that might be justified in terms of a child's right to an education. Notice that a right to education could be justified on grounds of the child's interest not only in future welfare but also in autonomy.

The remaining problem concerning rights and liberty is incorporating them systematically in schemes of moral evaluation. This problem takes different forms in different accounts of why liberty or rights are valuable. On instrumental or consequentialist views, the problem of rights articulation can be formulated as one of selecting rules that maximize good consequences. This is straightforward in principle, although difficult problems of strategic coordination may arise in devising effective rules. The choice of rules may also be influenced by limitations on persons' ability to calculate.

What about cases where rights protections are viewed as intrinsically valuable? One option here is simply to embed instances of rights protections (or of rights violations or losses of liberty) in the consequences, weighting them according to some scheme that reflects their importance relative to other aspects of outcomes. This approach seems to capture only part of the value of rights-protection. For in a situation in which committing one serious wrong, such as torturing a child, will prevent others from perpetrating two equally serious wrongs, this

⁴² Jules Coleman and Jody Kraus (1986, p. 31). See also Guido Calabresi and Douglas Melamed (1972). On this view, sometimes called the "choice" view of rights because of the emphasis on the need for the right-holder's voluntary consent to any invasion of the right, only what Calabresi and Melamed call "property rules" (which rest control with the right-holder) serve to define true legal rights. Coleman and Kraus argue persuasively that "liability rules," which address themselves to the right-holder's welfare rather than to the right-holder's control, also define legal rights.

approach would mandate carrying out the torture. But many would say that such actions are simply wrong, regardless of the consequences or, at any rate, that the moral pressure on me not to do this evil thing goes beyond an impersonal weighing of the consequences. The consequentialist view does not capture the sense that the moral imperative on me is that I should not torture a child.⁴³

One way to preserve a more thoroughly nonconsequentialist version of rights-protection is to adopt Robert Nozick's view of moral rights as absolute "side-constraints" on action.⁴⁴ On this view, an individual's obligation is not to strive for a minimum of rights violations, or a maximum of any social welfare function. Rather, agents may do as they wish, within the constraints imposed by rights. If everyone acts morally, then there will be no rights violations. It is not, however, permissible for anyone to violate rights in order to prevent other rights violations. Rights are simply not to be violated. This absolute lack of trade-offs not only of rights against other good things, but of lesser violations of rights against greater ones, is an implication of Nozick's view which has been found disturbing by many philosophers (for example Nagel 1983), and we suspect most economists would agree.

Amartya Sen has proposed another way to incorporate concern for moral rights into individuals' assessments of how they should act, which avoids the rigid repudiation of trade-offs in Nozick's

⁴³ It's more unsettling to note that I should in fact be willing to torture the child, on this consequentialist view, if doing so will prevent the torturing of one other child and will also produce some trivial moral gain, say raising somebody's utility by a tiny amount (Williams 1973). For a powerful defense of the importance of this nonconsequentialist aspect of judgments of right and wrong, see Nagel (1980).

⁴⁴ 1974, pp. 28–35. A moral view that emphasizes right action independent (or at least partly independent) of consequences is often called *deontological*.

view. This is to allow the evaluation of the goodness or badness of a state of affairs (or set of consequences) to vary with the perspective of the person doing the evaluating. In the torture example, an onlooker might readily conclude that it is morally better for one child rather than two to be tortured. But from the standpoint of the person who is doing the torturing, what is at stake is that person's own action—whether he or she tortures a child. One may act on one's own commitment to rights and judge that despite the good consequences of torturing a child, the act is wrong. Everyone might agree with both these judgments: from a neutral point of view, two tortures are worse than one; but from the point of view of one facing the prospect of torturing a child, carrying out the torture is morally worse than failing to prevent two tortures from happening. This account remains consequentialist but makes the evaluation of the consequences relative to the position of the evaluator (Sen 1982b, 1983a).

Some of these complications are probably unavoidable in any theory that takes rights seriously. However, for public policy purposes, the evaluator-relativity that Sen identifies is less likely to be a consideration than in personal morality, as the aim will be to appraise rights-systems from a suitably impersonal point of view. For this purpose, it may often prove adequate either to treat rights instrumentally or to embed their protection as an intrinsically valuable goal in an evaluator-neutral consequentialist moral framework.

3.1.3 *Equality*. An energetic debate about conceptions of equality has recently emerged that conforms roughly to the distinction between preference-based and more objective conceptions of well-being described in Section 3.1.1. Corresponding to the preference-based view of well-being are conceptions of

equality of welfare, while corresponding to more objective views is a conception (or a family of conceptions) of equality of resources.⁴⁵ The correspondence is rough, because welfare need not be identified with the satisfaction of preferences, and resources may be conceived of as the raw materials out of which individuals achieve a level of well-being rather than as a part of a more objective notion of well-being itself.⁴⁶

Despite these nuances, the intuitive contrast between equality of welfare and equality of resources is clear enough (Ronald Dworkin 1981a, 1981b). “Welfarist” views argue that what ultimately counts in moral evaluation are levels of welfare (usually understood as preference satisfaction), and therefore to the extent that there is concern with equality it should be concern for equalizing welfare. “Resourcist” views maintain that the levels of welfare people achieve depend on personal choices and characteristics that are no concern of social policy or public moral judgment. The proper concern of an egalitarian should not be welfare, but the resources people have available to pursue their own ends, whatever they happen to be.⁴⁷ “Resource egalitarianism” plainly appeals to some of the same intuitions as the somewhat vague notion of equal opportunity (see Arneson 1989). As resource and welfare

egalitarians can hold different theories of welfare, there are many distinguishable positions in the debate.

Part of the complexity stems from the need to analyze both external and internal resources. Equal distribution of resources such as rice or tractors can be achieved by dividing the commodities themselves up equally. The resulting distribution will generally not be efficient, but competitive trade from an equal division starting point might be taken to preserve equality of resources. In fact, when all resources are external and alienable, a competitive equilibrium arising in this way can be shown to be both Pareto optimal and “envy free” in the sense that no one will prefer another’s bundle of commodities to his own. Hal Varian (1974, 1975) has proposed the label “fair” for envy-free Pareto optima, and has defended this conception of fairness as ethically appealing (see also Duncan Foley 1967 and William Baumol 1986). An important stimulus to the study of envy-free allocations of resources has been their apparent ability to provide a standard of equity that does not involve interpersonal comparisons of welfare.

It is not so clear how to define equality of resources when some resources are internal. Consider differences in talent. Equal division of talent is plainly not possible. One might say that resources are equal whenever people have equal amounts of goods and equal amounts of leisure.⁴⁸ Unfortunately, if people’s tastes differ, this outcome will generally not be efficient, and trade from this equal division starting point may disrupt equality, in the sense that the competitive equilibrium resulting from trade will not be envy free. (For examples, a proof, and

⁴⁵ The interplay between conceptions of well-being and conceptions of equality has gone both ways. Sometimes conceptions of well-being are defended on the ground that they give more adequate conceptions of equality.

⁴⁶ Thus, Rawls’ conception of primary goods connects up with a resourcist view of equality, with primary goods as the resources. Sen’s capabilities view is like resourcist views in being concerned with objective features of persons’ existence rather than with their subjective satisfactions, but unlike more strictly resourcist views, it emphasizes outcomes (in the form of functionalities) and freedom to achieve outcomes, rather than inputs.

⁴⁷ Ronald Dworkin (1990) is an ambitious attempt to provide moral foundations for this kind of view as part of a defense of a conception of liberal equality.

⁴⁸ If everyone has the same tastes, this will be a fair distribution—Pareto optimal and envy free. See Varian (1974, 1985). (This assumes that problems of monitoring work effort to prevent shirking do not arise.)

further discussion, see David Schmeidler and Elisha Pazner 1974 and Varian 1985.) This difficulty has led economists to put forward two alternative definitions of equal resources, whose egalitarian properties will not be disrupted by trade.

The first of these is the notion of a "contribution-fair" distribution.⁴⁹ Imagine equalizing people's external resources but allowing them to retain the fruits of their differential talents. The resulting competitive equilibrium will not be envy-free, as those with more talent will tend to have consumption-leisure bundles that others envy. The equilibrium is, however, envy free over a space of consumption-output bundles. That is, no agent will want to trade places with another where trading places entails getting the other's consumption bundle but also producing as much as the other person. Contribution fairness essentially leaves talents or internal resources out of the package of things to be equalized, and accepts inequalities due to competitive rewards for talent as fair. This seems to fall short of what many people would think of as equality of resources.

Alternatively, one can define an "income-fair" distribution (Varian 1974, 1975; Schmeidler and Pazner 1974; Baumol 1986). The easiest way to think of this is to imagine everyone owning equal partial shares in the labor power (and hence the talents) of everyone. These shares and all external resources are divided equally. If one wants more leisure than the share of one's own labor power one happens to own (a very small share in a large society), one will have to buy it back from others. If we define a person's "full income" to include the market value of his or her leisure, the resulting competitive equilibrium will equalize full

incomes—hence the name "income-fair." Notice, however, that the leisure of the talented is a high-priced commodity. It will cost a talented person more to obtain any given level of leisure than it will cost others. Thus, if tastes are constant across the community, the talented, with equal full incomes and more expensive leisure, wind up worse off—equality of resources in this sense winds up overcompensating the untalented in terms of equality of welfare!⁵⁰

If resources can include internal qualities like talents, why stop there? As Roemer (1985a) notes, Ronald Dworkin argues for redistribution of the fruits of unequal talent, but finds inequalities resulting from unequal ambition acceptable. Yet differences in persons' willingness to expend effort must ultimately be traceable to some set of differences in their biological and personal histories, just as for talents. Why not treat the capacity to be ambitious as a resource? For that matter, some people may possess enzymes that allow them to extract more nutrition from a meal, or more satisfaction from a movie. Might we not think of these differences too as hidden resources? It is easy to see that, following this route, one might argue that any dif-

⁴⁹ This terminology is due to Baumol (1986). Varian (1974, 1975) has labeled the same notion "wealth-fair."

⁵⁰ Ronald Dworkin (1981b), aware of this problem, proposed instead to define equality of resources in terms of a hypothetical insurance scheme that would guard against the misfortune of low productive abilities. Dworkin argued that insurance would produce more palatable results than would the mutual ownership of talents, but economists have criticized his analysis of the hypothetical insurance market (Varian 1985; Roemer 1985a). Developing suggestions of Bruce Ackerman (1981, pp. 116, 132–33), Philippe van Parijs has defended a weaker egalitarian notion of "undominated diversity" (1989, pp. 50–54). Undominated diversity obtains if for each pair of individuals *A* and *B* there is at least one individual who prefers *A*'s total package of internal and external endowments to *B*'s and at least one individual who prefers *B*'s total endowments to *A*'s. Arguing from a commitment to undominated diversity and other moral principles, van Parijs has made a persistent case for an unconditional grant to be paid to every adult citizen (1989, 1990, 1991).

ference in achieved welfare levels resulting from apparently equal resources really is the product of hidden differences in resources.

Roemer (1985a, 1986a, 1986b, 1987) has proved a remarkable set of theorems showing that the only distribution meeting several reasonable-looking requirements for a resource-equalizing mechanism is equality of welfare. Roemer rightly presents this result not as a proof that equality of welfare is the only proper standard, but rather as an effort to clarify the underlying issues involved in distinguishing resourcist from welfarist views. He argues that these issues are of two kinds. One concerns the gap between ideal reasoning about just distribution and realistic limits on information. Equalizing welfare requires enormous amounts of information not only about resource distributions but also about what is going on in people's heads. When we add to these problems the incentives for misrepresentation created by any system for equalizing welfare, some simplified version of resource egalitarianism may be a suitable compromise.

The second issue is more philosophical. The distinction between those internal resources society may wish to compensate for and those other internal differences which it should not compensate for must turn on views about the nature of individual responsibility and about the nature of the person. We may wish to compensate people for being born blind, but not for being born without ambition. We may regard it as appropriate to compensate a person for being stuck with an affliction that brings constant pain, but not appropriate to compensate someone for being stuck with religious beliefs that induce him to starve himself. Articulating and defending these distinctions in a coherent way is a large project, which is far from being done (Scanlon 1975, 1986).

From this standpoint, relying on preferences even to the extent presupposed by the notion of envy-free allocations seems problematic. The whole project of deciding what counts as a resource and of deciding the relative weights to be given to different resources in assessing departures from equality demands more objectivity. The list of resources and appraisal of their relative weights need to be objects of public moral deliberation. As with Rawls' primary goods or Sen's capabilities, the valuation of these resources should depend on assessment of the importance of their role in facilitating a range of good lives. Evidence about the amount of preference satisfaction or the level of market value attaching to a particular resource should influence, but not determine judgments about the resource's value.

In addition to this rapidly developing literature addressed to the question of what to equalize, discussions of whether, why, and to what extent to equalize continue. Judgments about how much to equalize are plainly dependent on the role of equality in overall schemes of moral evaluation. A commitment to moral and political equality is deeply embedded in American political traditions and in Kantian moral philosophy, but the implications of that commitment for economic equality as such are controversial. Some theorists, including Ronald Dworkin and Rawls, argue in effect that the same fundamental commitments that rationalize support for moral or political equality should lead us to accept economic equality as a baseline standard, valuable in its own right and to be departed from only at a cost. For these theorists, underlying moral principles of impartiality and mutual respect carry implications for the degree of economic inequality that is morally acceptable. Other theorists, however, deny that economic equality as such is of any moral impor-

tance (Raz 1986, ch. 9; Frankfurt 1987). Notice that it is compatible with a skeptical view about the intrinsic moral value of economic equality to hold that it may have substantial instrumental value, on the grounds that economic inequality tends to give rise to other serious evils, such as widespread poverty, inequality in political power, and political or social instability. These instrumental questions give rise to empirical questions about, for example, the political influence of wealth and the influence of economic inequality on social stability, as well as to theoretical questions involved in the very separation of political and economic life. Although Michael Walzer, for example, argues in defense of a pluralistic society in which political equality and economic inequality coexist, he maintains that the authoritarian structure of economic corporations is political rather than economic and is as inconsistent with political equality as are towns that are owned and controlled by corporations (1983, ch. 12). Given the striking increase in economic inequality over the last decade and the evident influence of wealth on election results, these questions are bound to become an increasingly important part of the normative agenda of economists.

3.1.4 *Justice.* The last of the dimensions of evaluation we shall survey is justice, which some take as an umbrella term incorporating all dimensions of evaluation besides efficiency. But the focus here is on issues of justice apart from those touched on in the previous discussions of equality, rights, and liberty.

People in our society are committed to a large number of principles of justice with which they evaluate economic institutions and outcomes. Consider, as a first example, some of the many *prima facie* principles that influence whether employees regard a given wage structure as fair:

1. More difficult, unpleasant, or dangerous work should be better paid.
2. Jobs that carry a great deal of responsibility should be better paid.
3. Hard work should be rewarded.
4. Seniority is fair. Those who have worked for a given employer for a long time should be better paid.
5. Wages should not depend on race, sex, or arbitrary whim.

These are only a few of many such principles, and there are other principles establishing priorities in case of conflict among these. Moreover, not all of these principles link up easily to positions in moral philosophy, and economists and philosophers may want to dispute some of them or particular attempts to apply them.⁵¹ Some economists might question whether it is any more appropriate to offer a moral evaluation of wages than to offer a moral evaluation of the price of scissors or of the amount of snowfall in Buffalo, New York. But it is plain that moral judgments about wages on the lines enumerated above are frequently made by workers (if not by employers as well), and evaluation of such moral judgments may be relevant to the study of the weight and tenacity of the agents' commitments. A repudiation of this kind of moral evaluation itself rests on substantive principles of justice, and these require a defense.

Principles of justice influence evaluations of other aspects of economic life. Consider for example, equality of opportunity. While no one will dispute its importance, there is great disagreement on what equality of opportunity consists of. Is affirmative action implied by equality of opportunity or inconsistent with it?

⁵¹ For example, Joseph Carens (1985) argues that despite the intuitive appeal of the principle that dangerous and unpleasant work should be better paid, the principle is better served by reducing inequalities in wealth and income than by direct attention to relative wages.

Does equality of opportunity require that more resources be devoted to the education of the relatively disadvantaged? Does equality of opportunity require some more substantial egalitarianism, such as equality of resources? Questions about the extent to which economic policies contribute to equality of opportunity can scarcely be broached until the concept itself is clarified.

There is also a strong consensus in favor of providing a basic “safety net,” some minimal support to those who cannot support themselves. One might maintain that this consensus rests mainly on principles of benevolence, but Rawls has offered a powerful contractual argument for, in effect, the highest attainable social minimum. In his view (discussed further in Section 3.2.4 below), impartial rational individuals would choose to maximize the lifetime expectation of primary goods for the least well-off representative individual. If one takes income as a proxy for primary goods, this means that the income of the least well-off representative individual should be maximized. Although this distributive principle, called by Rawls “the difference principle,” has some objectionable features, Rawls has provided an important form of argument defending the justice and not merely the charitableness, of a social minimum.

For one last example, consider exploitation. What seems wrong about exploitation is that some people unfairly take advantage of others. But the traditional theories of economic exploitation have not made clear how the sort of exploitation they discuss involves anyone unfairly taking advantage of others. Marxian exploitation is reflected in the fact that wages are less than total revenue net of the costs of nonlabor inputs. But it has not been obvious why exploitation in this sense is a bad thing. The same difficulty applies to extensions of neoclassical theory such as Joan Robinson’s conception

of exploitation (1933, book IX) whereby owners of factors of production are exploited if they are paid less than the values of their factors’ marginal products.

Both in clarifying the traditional Marxian notions of exploitation and in offering an explication or replacement in terms of property rights, John Roemer has greatly clarified discussion of these issues. In his view capitalist exploitation (and there are other kinds) occurs if some group (the workers) would do better if they were to withdraw from society with their per capita share of the means of production, while some other group (the capitalists) would do worse.⁵² This view of exploitation, like the others, is open to the question, “So what’s wrong with it?” Roemer gives the striking answer, “It depends on how it comes about.” It is not capitalist exploitation itself, but the unequal initial distribution that is morally objectionable. If capitalist exploitation were to arise pristinely through the voluntary transactions of individuals whose initial holdings are morally legitimate, then there may be nothing wrong with it (1985b). In Roemer’s view a moral concern with exploitation leads one toward a historical view of justice, such as Nozick’s (Section 3.2.2 below) and away from a “patterned” view.

3.2 *Four Bases for Overall Moral Judgments*

Although treatments of detailed moral concepts and problems are, we believe, most likely to be of immediate use to economists, no survey could omit a discussion of the moral theories that economists and others use to guide their overall assessment of economic institutions and outcomes. Not only are there

⁵² 1982, ch. 7; 1988, ch. 9. This is a first approximation, and in his 1989 article Roemer accepts a suggestion of Erik Olin Wright that the greater benefits of the capitalists have to be linked to the effort exerted by the workers.

many specific moral theories, but they are of many kinds. To limit our survey, we shall focus on specimens of only four kinds of moral theories: (1) those which focus on efficiency, (2) libertarian theories, (3) utilitarian or consequentialist theories, and (4) contractarian theories.⁵³

3.2.1 *Efficiency and Pareto Optimality.* If one is minimally benevolent and favors making people better off, then, other things being equal, one should accept the principle (which John Broome calls “the principle of the personal good”) that “If two prospects are equally good for everyone, they are equally good, and if one of two prospects is better for someone than the other and at least as good for everyone, then it is better” (1989, p. 11). If in addition one identifies being better off with having a higher utility, then (other things being equal) one will endorse Pareto improvements and efficient (or Pareto optimal) allocations. Moreover, one can offer an argument in terms of these concepts for regarding a perfectly competitive equilibrium as a sort of moral ideal, other things being equal. Because competitive equilibria without externalities, public goods, informational limits, etc. are, by the first welfare theorem, Pareto efficient, they will be (other things being equal) morally good economic states, and market failures will be bad. Furthermore, the second welfare theorem shows that an efficient economic outcome producing any desired distribution of welfare can be attained by the workings of a competitive

market, given the right initial distribution of endowments to agents. So with no ethical premise more controversial than minimal beneficence, albeit with a controversial conception of individual well-being, economists can offer moral evaluations along one significant dimension.

However, the Pareto principle may be unappealing when it involves honoring people’s “nosy” preferences regarding others’ conduct (see the discussion of Sen’s liberty paradox below in Section 4.1). Moreover it has been shown that the Pareto principle is inconsistent with the conjunction of individual and collective rationality in conditions of uncertainty (Hammond 1983; Teddy Seidenfeld, Joseph Kadane, and Mark Schervish 1989). The problem lies in the fact that unanimity in individual rankings may rest on disagreements in subjective probability judgments and disagreements in preferences among options involving no uncertainty.⁵⁴

Quite apart from these problems, the Pareto principle or variants appropriate to different views of well-being have very limited applicability, as economic changes usually involve winners and losers. Economists have been reluctant to accept the severe limitations on efficiency judgments imposed by the Pareto criterion and have struggled for many years over the issue of whether some device could be found that would permit judgments in cases involving trade-offs among individuals without invoking interpersonal utility comparisons. The hope was that one could find a way to endorse “potential Pareto improvements,” in which the gains to the winners

⁵³ Recently there has been a good deal of striking criticism of liberal moral theorizing on the grounds that it illegitimately abstracts from the “rich” or “thick” moral discourse that characterizes communities. Authors such as Alastair MacIntyre (1981), Michael Sandel (1982), Charles Taylor (1979), and Walzer (1983) allege that moral theories of these kinds neglect or undervalue the values of community or public life, or that they presuppose a mistaken conception of the self abstracted from communal values and commitments. For a compact critical overview see Allen Buchanan (1989).

⁵⁴ For example, suppose the bourgeoisie and proletariat both prefer an armed worker’s revolt, because they form different estimates of its likely consequences. The fact of their agreement does not establish that an armed revolt would be socially better. See also Levi (1990).

were sufficient to compensate the losers in utility terms, *without actually making the compensation* (Hicks 1939; Nicholas Kaldor 1939). The conclusion has been that no such blanket endorsement is possible, partly owing to technical failings in the proposed devices for identifying improvements, but more fundamentally because there simply is no way to judge changes that affect distributions while remaining neutral on distributive questions (Little 1957; Streeten 1953; Samuelson 1950). But to stop here would leave welfare economists with virtually nothing to say, and practical applications of welfare economics in the form of cost-benefit analysis need to go on. The difficulties seem to have been dealt with in two somewhat unsatisfactory ways. One is by increasing modesty: one cannot always endorse potential Pareto improvements, but at least economists can identify them and leave the endorsing to politicians. Second, it seems that welfare economists have perhaps moved implicitly toward reuniting considerations of efficiency and equity. Blanket endorsements of potential Pareto improvements are unjustifiable, but as there are various theoretical ways to adjust such measures for equity concerns, one can suggest that equity-adjusted potential Pareto improvements can consistently be endorsed. But economists are queasy about the final step of actually endorsing any one way of adjusting for equity concerns.⁵⁵

3.2.2 Libertarianism. Libertarianism is both a political program and a philosophical position, and the latter is our main concern. The last two decades have seen a revival of philosophical libertarianism, “the doctrine that the only relevant consideration in political matters is individual liberty” (Jan Narveson 1988, p. 7; see also Tibor Machan 1982, p. vii).

⁵⁵ See also the discussion of “fair” allocations above pp. 697–98.

In libertarian writings, liberty—the fundamental value—is typically linked to a rights-based view of justice. Acts and policies are just if and only if they do not violate anyone’s rights. Moreover, the duties that rights imply are taken to be almost exclusively negative duties not to interfere. The right to life, for example, is a right not to be killed, not a right to be given subsistence. In holding such a view, libertarians are not necessarily endorsing selfishness, although some, such as Ayn Rand’s (1964) “objectivist” followers, are. Most libertarians value beneficence and charity as virtues. They only insist that individuals must not be forced to be beneficent or charitable. And some, such as Loren Lomasky, even defend some minimal rights to positive assistance, such as a right to be given a fair trial or an infant’s right to nurture (1987, pp. 164, 260). But Lomasky wants to limit such rights narrowly. The duties to provide these particular benefits are, like the duties not to interfere, easy to comply with without compromising the character of one’s projects, and the benefits are crucial to the possibility of pursuing any projects at all.

It is not obvious that taking liberty to be the fundamental value automatically commits one to a rights-based view of morality or to the view that rights rarely obligate others to give positive assistance. Indeed, Raz argues extensively that valuing autonomy commits one to an extensive role for government in providing the public goods that facilitate the achievement of autonomy (1986, ch. 8). But the arguments are complex and the libertarian case turns on the interpretation of the problematic notion of liberty.

In Robert Nozick’s well-known version of libertarianism, natural rights—rights that do not depend on consequences—secure individual autonomy. Justice is respecting rights. According to Nozick’s entitlement theory of justice, an outcome



is just if and only if it arises from just acquisition of what was unowned or by voluntary transfer of what was justly owned. Just acquisition is acquisition that violates no rights, and transfers are, in the relevant sense, voluntary just in case none of the limitations on individual choices arise from rights violations.⁵⁶ Justice is a matter of entitlement and depends on the actual history, not at all on the resultant pattern. Only the need to rectify past injustices justifies redistribution. Some of Nozick's most striking arguments in favor of this "entitlement view" involve attempts to show that even minimal efforts to regulate the patterns of distributive outcomes involve much more extensive interference in people's lives than is at first apparent. The libertarian system of rights is presented as having an appealing coherence and integrity. Because the point of rights for Nozick and most other libertarians is to secure liberty and to permit individuals to pursue their own projects, considerations of welfare never justify interferences with individual liberty. Nozick argues that only an extremely minimal state can be justified.⁵⁷

The (philosophical) libertarian's commitment to liberty is, in principle, independent of any consequences for human welfare, but libertarians would like it to be the case that protecting freedom also makes people better off, and libertarians need to show that some of their more extreme proposals, such as privatizing all streets and highways (Walter Block 1982) would not lead to disastrous consequences. Moreover, some theorists who are not libertarians in the strongly anti-consequentialist sense discussed above

⁵⁶ A view such as Nozick's thus places heavy demands on a theory of rights, which Nozick leaves largely undeveloped. See Nagel (1983).

⁵⁷ Notice that the libertarian political program does not follow automatically, as a more-than-minimal state might serve liberty better than a minimal state.

might be classified as libertarian on policy grounds. Friedrich Hayek, for example, argues that a broad range of important values, including economic prosperity, social innovation, and political democracy, are best advanced by keeping governmental action to a minimum (Hayek 1967, 1976). Most libertarian economists are probably less influenced by distinctive philosophical commitments than by specific reasons for doubting the efficacy of government intervention to improve efficiency or effect income redistribution.⁵⁸

3.2.3 *Utilitarianism and Consequentialism Reborn.* Consequentialism is the ethical theory that an action, rule, or policy is morally permissible or right if and only if there is no alternative with better consequences. Different varieties of consequentialism differ about what "better consequences" are. "Welfarist" moral theories take the view that only consequences for individual well-being ultimately matter, with other items such as rights or virtues viewed strictly as means to promoting welfare. Utilitarianism is, in turn, a subspecies of welfarism in which only the sum or average of individual welfare levels matter. The major 19th century utilitarians (especially Bentham, Mill, and Sidgwick) took utility to be happiness or pleasure.⁵⁹ But one can also regard utility, as economists do, as an index of preference or desire satisfaction.

⁵⁸ "[T]hose who toil in the libertarian fields devote the lion's share of their efforts to persuading us that the alleged benefits of the State are illusory . . ." (Narveson 1988, p. 183). For example, the literature on property rights and transaction costs that derives from Ronald Coase's arguments in "The Problem of Social Cost" (1960) usually reaches conclusions that are congenial to libertarians, but principled libertarians would emphatically reject the consequentialist view of rights as devices to reduce transactions costs and thereby to alleviate the suboptimalities caused by externalities.

⁵⁹ As Broome (1991a) points out, they actually took utility to be that property of objects that causes happiness or pleasure, but this complication is not germane to the discussion here.

During the decades that preceded the 1980s, utilitarianism appeared to be almost dead. Systematic moral philosophy had fallen from favor with the rise of logical positivism and its subjectivist and noncognitive view of ethics, and utilitarianism had the additional stigma of George E. Moore's (1903) critique of the so-called "naturalistic fallacy" (the claim that the good can be defined in "natural" or nonethical terms). Although utilitarianism continued to influence policy makers, most philosophers did not take it seriously as a moral philosophy. Even economists abandoned utilitarianism in the face of the difficulties posed by interpersonal utility comparisons. And to pound the final nails in the coffin, John Rawls, in *A Theory of Justice* (1971) offered not only systematic criticism, but an alternative theory that was suitable for guiding policy. The resurgence in practical moral philosophizing that was so prominent in the 1970s usually took for granted some sort of rights perspective rather than any sort of consequentialism.

Yet by the end of the 1980s utilitarianism and consequentialism were alive again and arguably dominant in both theoretical and applied moral philosophy. In the diverse works of authors such as Brandt (1979); Broome (1991b); Griffin (1986); Hardin (1988); Richard Hare (1981); Harsanyi (1977a); Kagan (1989); Parfit (1984); Peter Railton (1984); and Sen (1982b, 1983a), consequentialism has been resurrected and transformed. Although these authors defend very different ethical theories, with Brandt, for example, developing a sophisticated mental-state variant of utilitarianism and Sen developing a consequentialism in which rights, capabilities, and functionings play a more central part than happiness, there are common traits which make this work especially interesting and appealing to economists.

1. Most of these authors draw a specific connection between ethics and the theory of rationality. There has been a widespread recognition that refusing to make trade-offs among different objectives and to maximize some sort of objective opens one to charges of irrationality. In Samuel Scheffler's terminology, deontological (nonconsequentialist) ethical theories employ both "agent-centered prerogatives" (they sometimes *permit* agents not to act so as to maximize the good) and "agent-centered constraints" (they sometimes *prohibit* agents from acting so as to maximize the good). Although Scheffler argues that values such as integrity justify agent-centered prerogatives, which Kagan disputes at length, agent-centered constraints seem indefensible to both of them.

In any event the discussion has the nonconsequentialist on the defensive, and the charms of maximizing theories, to which economists have long since succumbed, have been seducing philosophers as well. Thus a good part of this new ethical theorizing is both more congenial to economists and more easily integrated into normative economics. None of this work is, to be sure, easily operationalized, but the difficulties of putting it to work in applied normative theorizing are not insuperable, either. Indeed some of this work, such as Harsanyi's, would in many cases justify returning to old-fashioned utilitarian policy analysis.

2. With the exception of Richard Brandt, no prominent theorist now defends a hedonistic conception of utility. All of the other specifically utilitarian theorists on the list join economists in taking utility not as an object of reference, but as an index of preference satisfaction. Yet none of them join economists in regarding utility to be an index of satisfaction of *actual* preferences. Although the refusal to identify welfare with the satisfac-

tion of actual preferences sets moral theorists off from most welfare economists, the distance is much smaller than in the case of rights theorists or hedonists, and there are greater opportunities for building bridges. It might, for example, be possible to operationalize some of the conditions moral theorists have placed on "rational" (Harsanyi) and "well-informed" (Griffin, Railton) preferences and thereby to strengthen the moral basis for the conclusions of welfare economics.

3. Some of this new work on utilitarianism and consequentialism is informed and influenced by developments in economics and game theory, and more of it could be. In Russell Hardin's utilitarianism, for example, human ignorance of consequences and the difficulties of measuring and comparing utilities hold center stage, and various concepts from both cooperative and noncooperative game theory are put to work in generating the outlines of a utilitarian theory of property rights and of their limits. But the game theory Hardin employs and his derivations are informal. There are important opportunities for collaboration between economists and philosophers in developing institutional implications of utilitarianism in circumstances of scarcity, uncertainty, opportunism, and so forth. We shall have a little more to say about these opportunities in Section 4 below.

4. Of particular interest in this resurgence of consequentialism has been the development of consequentialist moral theories that are not utilitarian—of theories that take the good consequences to be maximized to be other things than happiness or the satisfaction of preferences (esp. Parfit 1984). Although it is difficult to find empirical proxies for some of these good consequences, the more objective the goods to be maximized the more opportunity (other things being equal) for empirical application. If philosophers can specify a well-defined and

clearly measurable good to be achieved, then the welfare economist can step in and discuss how best to achieve it.

One important objective good in many of these theories is the satisfaction of needs, and even utilitarians emphasize it, though not, of course, as a fundamental and intrinsic good. For example, although James Griffin is a preference utilitarian, he argues that needs should have a great importance in social policy making on the grounds that governments can more easily tell what people need, than what will satisfy their informed preferences (1986, ch. 3). This emphasis on the empirical tractability of needs is somewhat ironic, given the traditional aversion economists have felt to distinguishing needs from mere wants. But this aversion seems not to have stemmed mainly from empirical difficulties in studying needs. It has arisen instead from qualms about the complexity that such a distinction would introduce into economic theory and from *theoretical* objections to drawing the distinction and giving needs any special weight. It seems to us that here is a case where moral philosophers are more practical than economists! In political discussions of economic policy, concern about human needs is already ubiquitous, and if philosophers can provide both a rationale for taking needs seriously in practical social decision making and a principled way of drawing the distinction between needs and wants in such contexts (David Braybrooke 1987; Garrett Thomson 1987), then economists can put their modeling tools to work in helping to devise policies that will satisfy needs.

Utilitarianism has, of course, been subject to many criticisms. Some of these, such as the view that it is a "pig philosophy," were unfair even when directed to traditional hedonistic versions, and do not apply to the preference utilitarianism that is now most common. Difficulties

concerning interpersonal utility comparisons continue, however, to plague utilitarians. Yet the most powerful objections to utilitarianism have always been intuitive. It is morally impermissible to murder one's nasty neighbor Max, even if it is obvious that total happiness would be greater without Max. Although these objections cannot be entirely defused, philosophers have done a great deal to undercut them. In his celebrated essay, "Two Concepts of Rules" (1955), Rawls pointed out that one must distinguish questions about the design of institutions from questions concerning the enforcement of institutional norms. Even if murdering Max would maximize utility, murder is forbidden by law and custom, and new laws and customs permitting murder whenever one conscientiously believed that doing so would maximize utility would not be favored by a utilitarian (or anybody else). Indeed Richard Hare (1981) argues that one should distinguish sharply between a critical level of moral thinking (which is typically beyond human capacities) in which, like an archangel, one considers all the consequences of alternative actions, and an intuitive level in which moral thinking is governed by rough principles to which people have strong emotional commitments. Hare argues persuasively that a utilitarian would want individuals to develop the sort of intuitions that would condemn even a utility-maximizing murder. Furthermore, as Hardin (1988) points out, such hypothetical examples usually presuppose knowledge that is simply unattainable. Utilitarianism remains unpalatable to most people (though perhaps not to most economists), but its difficulties are not obviously greater than those of alternative ethical theories.

Traditionally, most consequentialist ethical theories have offered very limited construals of the goodness of consequences. They have also imposed, im-

plicitly or explicitly, various kinds of neutrality constraints—for example, that the evaluation of a state of affairs should be independent of who is evaluating it or of who has brought the state of affairs about. Thus, for example, a utilitarian should not care whether Desdemona was run over by an ox-cart or stabbed by Othello, except insofar as the causal consequences of these ways of dying differ.

As we noted earlier, Amartya Sen (1982b, 1983a) has argued forcefully that there is no conceptual reason why consequentialist views have to be so abstemious. In principle, any feature that can be embedded in a description of a state of affairs can be an object of consequential moral evaluation. Consequences can include such items as rights violations (compare Section 3.1.2 above) and numbers and distribution of functionings achieved by members of society. There is similarly no conceptual difficulty in allowing such evaluations to vary among evaluators, so that Othello need not assess Desdemona's death from a neutral point of view.

One reason this broadening of the notion of consequentialism is important is that some prominent nonutilitarian moral theories have large consequentialist components. John Rawls' theory of justice, for example, is concerned with selecting that set of institutions which will maximize the achievement of equal basic liberty and, subject to that constraint, maximize the primary goods of the least well-off group in society—both of which are perfectly well-defined consequentialist goals.

It is important to distinguish the *conceptual* possibility of incorporating an extremely wide range of considerations in a consequentialist view from the *moral* question of what considerations belong in a moral evaluation. It is, for example, possible to define a consequentialist view

in which murders that happen on Tuesdays are morally permissible, unlike murders on other days. But it is morally unacceptable to permit evaluations to be sensitive to such arbitrary matters. The objections to deontological views noted by theorists as different as Kagan (1989) and Scheffler (1982) might well also be lodged against some extended forms of consequentialism.⁶⁰

3.2.4 *Contractualism*. “Contractualism” refers more to a method of arriving at moral judgments than to a substantive moral theory. The core idea of a contractual approach to morality is that an acceptable moral view for a society should, in some sense, reflect an agreement among the members of that society. The appeal of this core idea derives partly from the attractions of consent (James Buchanan 1975)—that persons should be bound only by what they have agreed to—and partly from an appeal to rationality—a morality that rational agents are willing to agree to under the appropriate circumstances is itself rational. It is useful to follow Barry (1989) in distinguishing two broad types of contractualist moral theory, especially as it applies to theories of justice. The first, exemplified by Thomas Hobbes and Hume, tends to identify rationality with self-interest and agreement with the outcome of bargaining, leading thus to a view of justice as *mutual advantage*. The second, exemplified by Kant and Rousseau, tends to identify rationality with the autonomous pursuit of ends (which need not be self-interested) and agreement with the outcome of a shared pursuit of common prin-

⁶⁰ Sen (1977b, 1985c, 1986b) has proposed that moral theories can be usefully classified and appraised in terms of the kinds of information they allow to count from a moral point of view. See also Charles Beitz (1986). On the particular question of whether different evaluators can legitimately come to different moral evaluations depending on their position see Sen (1982b, 1983a), Regan (1983), Kagan (1989), and Scheffler (1982).

ciples, leading thus to a view of justice as *impartiality*.⁶¹ Theories of both kinds have found much use for economic reasoning.

Thus, an early and influential attempt to formalize the notion of justice as impartiality came from Harsanyi's (1955) and Vickrey's (1945, 1960) independent derivations of utilitarianism. They captured the notion of impartiality by asking what moral principles rational agents would agree to (a) if they were deprived of knowledge of which people in society they were and (b) if (as a consequence of applying the principle of indifference) they believed they were equally likely to be anybody. People in this position, it was argued, would maximize expected utility by adopting the utilitarian principle for society as a whole: maximizing the average welfare of its members.⁶² Such a contractual approach leads naturally to average utilitarianism—maximizing average welfare—rather than classical utilitarianism—maximizing total welfare.⁶³ The exercise involves deciding how self-interested people would choose in this hypothetical situation, but the decision to conform one's conduct to the resulting principles in real life is *not* to be explained by self-interest: there is no

⁶¹ The differences are not perfectly sharp, as those who see justice as mutual advantage also see their principles of justice as impartial, while those who regard justice as impartiality also see adherence to their principles as mutually advantageous. But the difference in emphasis is important enough to be regarded as a difference in kind.

⁶² Though one might well wonder, as Gauthier (1982) does, how agents are supposed to maximize expected utility without knowing their preferences.

⁶³ These versions of utilitarianism will differ in their analysis of optimal population—with total utilitarians wanting to increase population until the marginal utility of another life is zero, while average utilitarians stop where the average utility of a life is maximized. The classical version is more naturally gotten not from a contractual argument but the perspective of a benevolent impartial spectator who sees utility, pleasure, or desire satisfaction of individuals as the only intrinsically valuable things in the world, and therefore seeks to maximize them.

claim that *my* interest is always best served by maximizing the *average* utility.⁶⁴

Since the classic articles by Harsanyi and Vickrey, various refinements of and objections to their stipulation of the initial situation in which impartial agreement is to be reached have been put forward. The most famous has been that of John Rawls. Rawls argued that rational actors behind a “veil of ignorance” about their individual identities would not choose utilitarianism. Rawls specifies the conditions of the initial situation—and the character of the actors in that situation—in more detail than Harsanyi and Vickrey do, and he makes two especially important modifications. First, Rawls suggests that agents behind the veil of ignorance will choose among principles that determine their life chances on the basis of which will best secure the basic means (primary goods) required to live out any of a number of rational life plans. Not knowing who they are nor what their preferences are, they are not and cannot be utility maximizers. Second, Rawls argues that rational actors in what he styles “the original position” will not rely on the principle of indifference, thus further shortcircuiting the argument for utilitarianism.

Rather than electing to maximize utility, claims Rawls, actors subjected to the radical uncertainty of the original position, and aware of their fundamental in-

terest in securing the means to a decent life, would adopt a “safety first” strategy in agreeing to principles of justice. They would seek first to safeguard their basic political and personal liberties and then, subject to that constraint, they would judge basic social and economic institutions according to their tendency to promote the interests of those least well off in society. The motivation for this maximin outlook is that it guards against the unhappy accident of being a member of that disadvantaged group.

In their full elaboration, Rawls’ principles of justice are:

First, each person is to have an equal right to the most extensive basic liberty compatible with a similar liberty for others. (1971, p. 60)

[Second] Social and economic inequalities are to be arranged so that they are both (a) to the greatest benefit of the least advantaged and (b) attached to offices and positions open to all under conditions of fair equality of opportunity. (1971, p. 83)

In addition, Rawls specifies that the first principle is, in social conditions such as ours, lexically prior to the second. These principles pose questions concerning the relations between liberties and economic welfare, concerning the role of market vs. nonmarket institutions in allocating goods, and concerning savings for future generations. The “priority of liberty,” as the first principle is sometimes labeled, is not easy to capture in economic models, and the lexicographic ranking of this principle is not the sort of thing economists find easy to swallow (but see Cooter 1989). However, part (a) of the second principle—that society should maximize the well-being of the least advantaged—has an elegant formal interpretation as a “maximin” standard. It implies that (subject to the first principle and part (b) of the second principle) society should maximize the primary goods of the least well-off. Although a focus on only this dimension of Rawls’ theory yields a somewhat

⁶⁴ Scepticism about interpersonal comparisons might lead a contractualist to stop short of average utilitarianism, instead arguing that a rational person in this situation would at least endorse the Pareto principle. Some have tried to squeeze a defense of the “compensation principle” (roughly, attain all potential Pareto improvements) on the (empirically dubious) grounds that in the long run gains and losses from individual potential Pareto improvements will tend to even out, so that everyone would rationally agree to all such changes, in the absence of specific knowledge of his or her own situation (Richard Posner 1979). See also Streeten (1953) for discussion of such views.

distorted picture of Rawls' actual view, this "maximin" interpretation of Rawlsian justice has led to some simple and valuable comparisons between Rawlsian and utilitarian analyses of ideal economic arrangements (for example, Edmund Phelps 1973; Sidney Alexander 1974; and Richard Musgrave 1974).

Both utilitarianism and Rawlsian justice as fairness provide substantial scope for redistributive activity by governments to raise average welfare or to protect the least well off. This result is characteristic of theories that emphasize justice as impartiality. Notice, again, that the argument is not that redistributive taxation is in the interest of those who must pay the taxes, but rather that taxpayers should acknowledge that, looked at impartially, it is fair that they should pay. What motivates people to support such taxation is not, on this kind of account, rational self-interest, but some form of moral concern.⁶⁵

Theories of justice as mutual advantage give self-interest a larger role. The rules of justice are conceived as the result of rational bargaining among well-informed and self-interested actors. Gauthier, for example, argues that if we suppose that such agents lack the means to make interpersonal utility comparisons, then they would agree to distribute the gains from cooperation in accordance with a principle of "minimax relative concession."⁶⁶

Gauthier argues that this principle would fairly distribute the gains from social interaction relative to the situation that

would prevail in the absence of agreement. As, in Gauthier's view, this non-agreement point would itself reflect substantial inequalities, which are largely preserved in the just society (else it would not be in everyone's interest), the scope for redistributive governmental activities is substantially less than on Rawlsian or utilitarian views.

Gauthier, like others who want to characterize justice as the outcome of a mutually self-interest bargain, wishes to avoid reliance on an independent moral motivation in explaining people's willingness to comply with justice. The problem is a challenging one, as the presence of rules of justice creates a classic prisoners' dilemma: even if you and I jointly benefit from the rules, I will do still better if you obey them and I violate them when it suits me. Gauthier argues that, under certain circumstances, it will be in my rational self-interest to cultivate the kind of character that will lead me to obey the rules of justice, even if having that kind of character sometimes robs me of opportunities to free-ride. It is crucial to this solution that other members of society be able to detect, often enough, what kind of character I have, and to reward or punish me suitably, in terms of their willingness to interact with me, depending on what they discern. The issue here—of when it may be self-interestedly rational to cultivate dispositions that interfere with acting on self-interest—has received valuable treatment from economists (Akerlof 1983 and Frank 1988).

From a somewhat different perspective, James Buchanan and his collaborators argue that, as an empirical matter, self-interested agents who are concerned about "constitutional" choices—choices involved in setting the general rules of the game—face so many uncertainties that they might as well be behind Rawls' veil of ignorance. Because the gains from constitutional design may be very large,

⁶⁵ Scanlon suggests that the basic motivation for complying with contractualist moral principles is a desire to "justify one's actions to others on grounds that they could not reasonably reject—reasonably, that is, given the desire to find principles which others similarly motivated could not reasonably reject" (1982, p. 116).

⁶⁶ This is the Kalai and Smorodinsky bargaining solution, which is discussed below, page 722. In most circumstances minimax relative concession coincides with equal relative concession.

self-interested agents should be concerned about such problems, and they thus have reason to step behind “the veil of uncertainty” (as Geoffrey Brennan and James Buchanan 1985 style it). But in Buchanan’s view rational agents behind the veil of uncertainty will be much more cautious about redistribution, because they will model both economic and political behavior as self-interested. They will make this assumption partly on grounds of realism and partly on the Humean grounds that in designing constitutions “every man ought to be supposed a knave” (quoted in Brennan and J. Buchanan 1985, p. 59). Thus, even supposing that redistribution of market incomes toward the poor is desirable, it is not obvious that political institutions such as majoritarian democracy will produce such redistribution (Brennan and J. Buchanan 1985). Given the vagaries of politics, individuals adopting a constitutional perspective would not want to leave redistribution to the discretion of self-interested politicians. If they favor equalizing measures at all, they may prefer to embed particular distributive rules, say a negative income tax financed by a flat rate tax, directly into the constitution. Such constitutional distributive rules should not in fact be considered redistributive, as they would figure into the underlying definition of property rights.

The general idea that moral principles or social norms can be usefully interpreted as the outcome of self-interested bargains is receiving wide attention. Along with Gauthier and James Buchanan, Jean Hampton (1986) and Gregory Kavka (1986) have outlined general theories of justice that adopt this Hobbesian starting point. The same general idea has also been employed to consider what sorts of detailed rules or conventions would become established among agents in various relations to one another. Thus Andrew Schotter (1981),

Sugden (1986), and Michael Taylor (1987) consider the recurrent problems of coordination and conflict individuals face and the sorts of institutions and conventions they will come to accept to resolve them. Ken Binmore (1990) argues that imperfectly rational agents will be led to a utilitarian solution to recurrent bargaining problems. This sort of disaggregative contractualism places heavy demands on the tools of game theory (see Section 4.2 below). Its findings may have both explanatory and normative force.

Even economists who are not interested in the foundational questions raised by contractarian theories of justice are likely to be interested in the implications for economic policy and institutions of competing contractarian views. There has been considerable interest in recent years in working out the implications for tax policy of alternative versions of both Rawlsian and utilitarian distributive principles (James Mirrlees 1976; Phelps 1973; Atkinson and Joseph Stiglitz 1980, chs. 12, 13, and 14). The brand of contractarian argument advanced by James Buchanan (and having its origin in J. Buchanan and Gordon Tullock 1962) has encouraged analytical work in more of a “political economy” vein, involving efforts to work out the equilibrium behavior of alternative constitutional regimes under the assumption that actors are self-interested in both politics and economics (see Brennan and J. Buchanan 1985, ch. 8).

3.3 *Concluding Words on Evaluation*

Few economists can avoid ever making appraisals of economic outcomes or institutions, and in making such appraisals they unavoidably rely on some moral philosophy. The most obvious strategy for economists to employ is to rely on the most robust and least controversial moral premises that they can. For many purposes this strategy is a good one, al-

though economists' perceptions of what is uncontroversial and what is not may be skewed by their theoretical commitments. But for those economists with a deep concern with evaluation, this policy of laying as few hostages as possible to the vicissitudes of arguments in moral philosophy is bound to be unsatisfying. Very little in ethics is completely uncontroversial, and very little can be said about economics that relies on only uncontroversial moral premises. Economists concerned with evaluation are going to have to get their feet wet in the swamps of moral philosophy. We've tried to provide a rough map of the terrain, but we make no guarantees that those who venture in will not find some surprises. It's tough and tangled territory.

4. *Technical Aspects of Applied Moral Theory*

Over the last four decades economists and decision theorists have made exciting progress in the development of formal characterizations of individual and social choice, strategic interactions, and related concepts. A good deal of this work seems to be linked to moral philosophy. Formal models of rationality and game theoretic studies of incentives hold out the hope of superseding the ancient puzzles concerning the relations between morality, rationality, and self-interest. Concepts of "fair" or envy-free allocations, discussed in Section 3.1.3 above, facilitate the articulation and evaluation of egalitarianism. Solution concepts in game theory may enrich the contractarian perspective that morality arises from or can be justified in terms of agreement. Formal work in social choice theory tests the consistency of traditional principles concerning how social policies should respond to individual interests and respect individual rights. We face an embarrassment of riches and can only offer a highly selec-

tive survey of relations between a few of these developments and moral philosophy.

4.1 *Social Choice Theory*

Social choice theory grew out of Abram Bergson's (1938) and Samuelson's (1947) formulations of the "social welfare function," and was given its decisive formulation by Arrow in his *Social Choice and Individual Values* (1951). Social choice theory has had a major impact on the way economists conceive normative problems and on contemporary moral philosophy as well. However, the interpretation of theorems in social choice, and even the definition of the subject, have been elusive and controversial.

As Sen (1986a) has suggested, social choice theory might be defined either by its approaches or by its subject matter. Sen defines its subject matter as follows: "Social choice problems arise in aggregating the interests, or preferences, or judgments, or views, of different persons (or groups) in a particular society" (1986a, p. 214). Because aggregating interests and resolving conflicts is central to moral philosophy, a large part of the subject matter of social choice theory is moral philosophy. But the methods of social choice theory are distinctively formal and axiomatic. Most works in social choice theory are devoted to proving theorems.

Although one might usefully regard social choice theory, as Sen suggests, as the union of its subject matter and approach, what most people think of as social choice theory is the intersection: the proof and interpretation of theorems concerning the social aggregation of preferences, judgments, and interests. The relevance of social choice theorems to morality depends on exactly *which* attitudes or judgments are being aggregated, and *for what purpose*. Sen has hopefully distinguished two broad purposes of aggregation, namely *deciding*

and *evaluating*. Thus, one might want an aggregation mechanism for deciding which movie a group of people was going to see, and also an aggregation mechanism for determining which among a set of movies the group has seen was best, from an aesthetic, moral, political, or comic point of view. The criteria for judging whether a mechanism was good from the first point of view might be very different from those that should govern in the latter context.⁶⁷ For example, it seems more imperative for social evaluations than for social decisions to be transitive, while the Pareto principle is more compelling for social decisions (Broome 1989).

As the movie example suggests, *what is being aggregated also differs among cases*. Sen distinguishes among “interests, preferences, judgments, and views” as objects of aggregation exercises. No doubt this list could be extended to other items including capabilities, resource endowments, and so on. Any of these properties of individuals might be aggregated either for purposes of arriving at a social decision or an evaluative judgment. A good deal of the social choice literature treats the items to be aggregated as preferences.

Insofar as social choice theory attempts to provide a social evaluation based on profiles of individual welfare, it is a branch of moral philosophy, and its theorems can reveal to moral philosophers ambiguities or hidden implications of accepted principles. A simple illustration of an interest-aggregation mechanism based on a normative principle is a utilitarian social welfare function, which de-

finer the socially best outcome as that which maximizes the sum of individual utilities (where the appropriate measure of individual utility may differ among different versions of utilitarianism). It is important to be clear on the point that what is involved here is not an attempt to aggregate individual normative judgments about what is socially best into a social judgment; rather utilitarianism presupposes the judgment that the best state is that which maximizes aggregate utility.⁶⁸

Moral reasoning has crucial roles in appraising particular social choice theorems. The most obvious role involves the assessment of the axioms. Do they express plausible normative principles? But moral presuppositions can also be embodied in subtler features of a formal setup. For example, treating the relation between profiles of individual preferences and social choices as a *function* implies the uniqueness of social choice. Furthermore, one’s moral scrutiny need not be limited to individual features of the formal system. One can question whether a particular aggregation mechanism as a whole is a morally justifiable way of arriving at either a decision or an evaluative judgment. One may, for example, object to a process for aggregating interests for purposes of decision making as *unfair*, perhaps because of the way it responds to minority interests. And one may object to a process for aggregating individual normative judgments into a collective judgment because it is *unwise*, in that it fails to respond to the quality of the arguments supporting a particular judgment. Finally, moral reasoning may also enter into the scru-

⁶⁷ Decisions and evaluative judgments are obviously related, but they are identical only on the strong assumption that the operative decision rule is always to choose the best alternative according to the given evaluative standard. But there can be more than one evaluative standard, and decision rules may have procedural virtues.

⁶⁸ What if individuals get satisfaction from having their view of justice or of moral goodness prevail? Coherent and plausible versions of utilitarianism are unlikely to allow the measure of individual utility that goes into the social welfare function to be defined so broadly.

tiny of the individual objects that are fed into the aggregation mechanism. We reviewed earlier the likelihood that individual preferences may need to be “laundered” before becoming suitable objects to be accorded moral weight, and some similar process might be applied on moral grounds to interests, views, or judgments. One might, for example, want to make sure that the informational basis of these items was adequate.

If the axioms and formal set-up pass preliminary scrutiny, one can reverse the order of dependence and argue from the results of social choice theory to moral conclusions. Proofs are, after all, arguments. And if the axioms of a particular social theory are morally correct and the proof is valid, then the theorem must be morally correct as well. Axiomatic arguments of this kind have been proposed for ethical conclusions as different as utilitarianism (Fleming 1952; Harsanyi 1955; Vickrey 1945, 1960), “leximin” utility (Steven Strasnick 1976; d’Aspremont and Gevers 1977; Hammond 1976b), and welfare egalitarianism (Roemer 1988). Harsanyi, for example, demonstrates that if (1) personal and moral preferences satisfy the axioms of expected utility theory, (2) the preferences of different people are treated symmetrically, and (3) the Pareto principle is satisfied, then social choice must be utilitarian.⁶⁹

Although a good deal of social choice theory seems to consist of formal arguments for moral principles, we doubt that

⁶⁹ For a recent challenge to the cogency of this interpretation of Harsanyi’s theorem, see John Weymark (1991). Among other difficulties, the cardinal representation theorem (p. 681 above) only says that preferences satisfying the axioms *may* be represented by a cardinal utility function, U , with the expected utility property. The preferences may also—and apparently equally well—be represented by a utility function such as U^3 , which does not possess the expected utility property. Harsanyi needs to explain why the representation that yields cardinal expected utilities is the right one to use for ethical comparisons.

this is the best way to appreciate this work. Logically valid arguments only show that one cannot both accept all their premises and reject any of their conclusions. If one is convinced of the correctness of the premises, they can persuade one to accept the conclusions, but it is equally true that if one is convinced of the incorrectness of the conclusions, one will have reason to doubt the premises. As Kotaro Suzumura (1983, p. 3) quotes Samuelson (1977, p. 85): “the classic phrase ‘by their fruit must ye know them’ applies above all to axiom systems.” In general, our confidence in the axioms will derive both from how reasonable they seem to be in their own right and from our sense of the reasonableness and moral attractiveness of the theorems they lead to. If the axioms are shown to entail a moral conclusion that we find disturbing, we will be motivated to scrutinize them more closely, as well as to reflect on whether our negative reaction is really justified.

We suspect that the most important role of purported demonstrations of moral principles by social choice theorists has been to reveal the ambiguities and difficulties in apparently plausible moral principles. These ambiguities and difficulties have been highlighted even more conclusively by demonstrations that sets of plausible principles may be logically inconsistent. Arrow’s General Possibility Theorem is the most famous of these demonstrations. Arrow showed that there is no possible method of aggregating individual rankings of social alternatives that meets certain apparently innocuous criteria. Exactly what moral lessons should be drawn from this celebrated theorem is not however easy to say.

Much depends on how the individual rankings are interpreted. One possibility is to take the rankings as showing how social alternatives affect individual wel-

fare levels. Then what Arrow shows is that individual ordinal rankings cannot provide an acceptable basis for reconciling conflicts among individual interests or for making decisions about how welfare should be distributed. But as Sen has observed, this failure is not particularly surprising, because individual ordinal rankings provide little information concerning the relative importance of different interests. If interpersonal comparisons of preference intensity are permitted, or if nonpreference information (say on comparative income or resource levels) is permitted, the impossibility result may no longer hold (Sen 1970a, chs. 7, 8).

A different interpretation is that individual rankings reflect individual judgments about what is best for society to do. We might take these to be individuals' moral judgments over social alternatives. It seems less appropriate than in the case of interest aggregation to avoid the Arrow problem by introducing some kind of interpersonal comparison—in particular, it is not clear that the intensity with which a person holds a certain view should necessarily add to its weight in social judgment or decision making. What the Arrow theorem then tells us is that no acceptable scheme for aggregating these individual moral judgments will produce a consistent set of judgments for the community as a whole. Here again the result is not really surprising, when it is remembered that one of Arrow's conditions is that an aggregation scheme must work for all possible preference profiles. The theorem then says that no aggregation scheme can produce a consistent set of moral judgments over the social alternatives facing a society, no matter what individuals' moral judgments are. The Arrow conditions do not allow any sets of moral rankings to be ruled out as improper, and they do not allow for any modification of individual

moral views in response to discussion.⁷⁰ When constraints such as "single-peakedness" are imposed on preference profiles, the impossibility results no longer hold. Arrow's general possibility theorem does not, then, undercut such familiar practices in making moral judgments on public matters as weighing the importance of competing interests or deliberating over the persuasiveness of particular people's moral judgments and arguments (and Arrow has never suggested otherwise).

Sen's remarkable "Paretian liberal" paradox arguably has even more bearing on issues of social morality than Arrow's impossibility result. Sen showed that an apparently weak necessary condition of respecting individual liberty—that two individuals each be decisive over their own single pair of alternatives, which differ only with respect to some purely private matter—is incompatible with the achievement of Pareto optimality when people hold certain plausible rankings of social alternatives (Sen 1970b, 1976a, 1983b, and 1986a). His motivating example is one in which each of two people is acknowledged to have the right to read or not to read any given book and the two have conflicting preferences about what the other should do in regard to an erotic novel. The lewd would most like that both he and his counterpart, the prude, read *Lady Chatterley's Lover*. If only one should read it, however, Lewd wishes it to be his prudish friend, because it would loosen him up. Prude would best prefer the book not be read at all, but would rather read it himself than further his friend's corruption. Plainly, the Pareto-preferred solution is for Prude to read the book, but the exer-

⁷⁰ For criticisms of this conception of politics and advocacy of more deliberative views of democratic decision, see Joshua Cohen (1986, 1989), Jules Coleman and John Ferejohn (1986), Elster (1986), and David Estlund (1990).

ise of rights will result in Lewd reading it.

Allan Gibbard pointed out that if one strengthens Sen's minimal liberty condition to permit two individuals to be decisive over all pairs of alternatives that only differ with respect to purely personal matters, then certain combinations of preferences will rule out the possibility of any social choice, regardless of questions of Pareto optimality. In Gibbard's example of the nonconformist and Mrs. Grundy (1974, p. 389), the nonconformist wants his bedroom walls to be a different color from Mrs. Grundy's, while she wants her bedroom walls to be the same color as the nonconformist's. If each is decisive over social alternatives that differ only with regard to the color of his or her bedroom walls, every possible outcome will be vetoed by somebody.

In interpreting the significance of Sen's result, and of the extensive response to it, it is worth noting that both the example and the theorem are aimed only at one type of liberty, namely personal liberty over "private" matters (Jonathan Riley 1989, 1990). One essential feature that generates the paradox is then the fact that people are assumed to have preferences over outcomes that properly—according to the liberty principle—belong in someone else's protected sphere. Indeed, Julian Blau (1975) shows that Sen's paradox does not arise in the two-person case unless at least one person has "meddlesome" preferences in a technically defined sense. The Pareto principle, as applied in this context, treats these preferences on a par with personally oriented ones.

A number of responses to the paradox have attempted to dissolve it by introducing devices that modify the libertarian rights so as to preserve the Pareto principle in the face of these "nosy" preferences. Blau (1975) has proposed, for example, that libertarian rights should only

hold for a particular person when others' preferences over the outcomes within that person's "protected sphere" are not too strong. Gibbard (1974), Nozick (1974), Peter Gardenfors (1981), Sugden (1985), and Wulf Gaertner, Prasanta Pattanaik, and Suzumura (forthcoming) have all argued that the paradox arises from a misconstrual of rights. To have a privacy right is to have a certain freedom to do as one pleases, not to be decisive over social choices. Rights should be modeled as "game forms" (Gibbard 1974) rather than as a constraint on social choices. I have a right to determine the color of my walls if certain strategies are open to me. The fact that Mrs. Grundy will copy me is no violation of my rights. The game-form construal of rights would not only dissolve Gibbard's paradox, but it would permit individuals to enter into trades or contracts that allow them to arrive at Pareto optima. In the Prude/Lewd example, Prude might contract to read the book on condition that Lewd doesn't. Such optima may not be reached because of imperfect information and transaction costs, but the general incompatibility between liberty and efficiency would dissolve.

But there would be serious enforcement problems, for as Riley has shown (1989; see also Kaushik Basu 1984), in games corresponding to the social choice problems in which there is a clash between rights and Pareto optimality, the outcomes that respect rights will not be Nash equilibria (see footnotes 73 and 77 below). As Sen notes, there are questions whether such contracts or their enforcement is morally desirable, particularly because enforcement would require close supervision of what are agreed to be private matters. And, as Sen (unpublished) argues, the game-form interpretation of rights may be too weak. It is hard to see how any system of rights that provided genuine protection to individual

liberty could fail to satisfy Sen's minimal liberty condition—to make at least two individuals decisive over a single separate pair of social alternatives. The moral of the story in Sen's view is that normative social policy cannot be based only on preferences.

The other obvious direction to look for resolution of the liberal paradox is through relaxation of the Pareto principle. There is no evident moral requirement that preferences regarding other people's private activities should be accorded the same moral importance as those regarding one's own activities or public matters.⁷¹ Indeed, acknowledgement of a private sphere of activity for a person creates a presumption that those preferences should receive less weight. Other moral arguments have been offered for discounting other-regarding preferences in making public decisions (Ronald Dworkin 1977, ch. 12). It does not follow from these doubts about the Pareto principle that nosy preferences should be ignored entirely, nor that a simple formula can determine the weight they should receive in every case. But Sen's paradox does underscore the importance of permitting other things besides preferences to count. In particular, one may want to examine the motivation behind preferences in deciding on the role they should play in social decisions and moral judgments—a point we have noticed elsewhere in this essay.

What do examples such as the Arrow and Sen paradoxes, and the proofs of alternative moral systems from sets of axioms tell us about the fruitfulness of social choice theory in illuminating moral issues? While its subject matter—the problem of assembling individual judgments or interests into social evaluations—is unavoidable, the axiomatic ap-

proach which has characterized most work on social choice theory leaves us more ambivalent.

It is clear that the precision of formulation and the rigor in argument induced by the formal approach can lead to important clarifications, and can stimulate a new awareness of the depth of problems—as the literature on the liberal paradox shows. At the same time, the abstractness of these methods has pitfalls. To quote Sen: “Arrow was undoubtedly right in saying that ‘one of the great advantages of abstract postulational methods is the fact that the same system may be given several different interpretations, permitting a considerable saving of time’ (Arrow 1951, p. 87). But *that* probably is *also* one of the great disadvantages of these methods” (Sen 1977c, p. 190). Considerable confusion has been generated, particularly in discussions of the Arrow theorem, by a failure to distinguish among alternative interpretations of both the *objects* being aggregated and the *purposes* of aggregation. It is difficult to spend much time with the social choice literature without becoming aware of the disparity between the large amount of attention lavished on the derivation of theorems and the small amount devoted to interpretation of the assumptions and of the point of the exercises. Unless guided by ethical sophistication, social choice theory can degenerate into empty formal exercises. But it would be unfair to single out social choice theorists for blame. Moral philosophers ought to be more involved with this work, too. For if formalizations of plausible moral principles lead to inconsistencies or to unacceptable conclusions, then those moral principles may need revamping.

4.2 Game Theory

Game theory is concerned with strategic interactions among individuals who

⁷¹ For other grounds to question the ethical appeal of the Pareto principle, see Section 3.2.1 above.

are to some extent rational. An interaction among individuals is strategic if the choices of some individuals depend on the choices of others. Most game theory employs standard conceptions of rationality.

There is almost an unlimited variety of strategic environments in which individuals can interact. To make the problems formally tractable, game theorists typically suppose that the only relevant properties of individual strategies⁷² are (a) their causal consequences for the plays made by others and (b) the utilities of the outcomes of different combinations of strategies—the “payoff structure.” In game theory the players come with a knowledge of what the possible strategies are and at least a probability distribution over the possible utility consequences. Much of game theory also includes a “common knowledge” assumption—that players know what others know, including that others know that they have this knowledge, and so on.

Most game theory is not addressed to moral questions, and moral philosophy is less directly relevant to the scrutiny of particular axiom systems than it was in social choice theory. But moral questions concerning prudence, benevolence, and justice remain relevant. Not only does most game theory specify the values and choices of agents in terms of preference satisfaction, but game theoretic analyses of social interactions carry an implicit commitment to a highly contestable set of answers to some of the deepest questions in normative social and political philosophy—to the view that the only (and hence the proper) perspective for individuals to adopt in their social interactions is individual maximization. This perspective does not rule out

altruism or sympathy, and it need not in principle rule out “commitment” (p. 688 above), but it does rule out a collective perspective, a perspective that considers what *we* should do and what the consequences will be for *us*. And it has been argued that such a collective perspective is essential to the very notion of social interaction (Margaret Gilbert 1990) and to the explanation for why real individuals play simple games so differently than game theory would prescribe and predict (Robyn Dawes, Alphons van de Kragt, and John Orbell 1990; see also Hurley 1989, p. 145 and Regan 1980, ch. 8).

The relevance of game theory to moral philosophy is most obvious in normative applications of game theory to determine the “right” or “best” solution to problems of strategic interaction.⁷³ But positive applications of game theory to determine how more or less rational individuals will solve strategic problems may also be morally relevant. Because contractarians believe that what is morally right is what rational agents would agree to under appropriate circumstances, and game theory may determine the content of such agreements, game theory may have a role in determining what is morally right. We shall shortly see how David Gauthier employs game theory to argue for principles of justice, and as mentioned before (p. 711) game theory is central in attempts to see specific moral principles as solutions to detailed interaction problems. Game theory also bears on the question of the compatibility between the dictates

⁷² A *strategy* is a complete recipe for how a player will respond to any possible sequence of eventualities in the game.

⁷³ Work in game theory is also a crucial adjunct to social choice theory, for even if there are solutions to the problems of aggregation with which social choice theory is concerned, there are important questions to be asked about how such solutions can be attained (Dasgupta, Hammond, and Eric Maskin 1979; Bezalel Peleg 1984). Indeed, as we noted briefly above in discussing Sen’s paradox, many of the aggregation problems translate into difficulties concerning the solutions of games (Maskin 1985).

of individual rationality and the demands of morality.

Human flourishing requires social cooperation, but, as philosophers have long recognized, such cooperation is problematic. In Hobbes' state of nature, where there is little cooperation, "the life of man [is] solitary, poor, nasty, brutish, and short" (1651, p. 107). The theory of games is particularly relevant to moral philosophy in its revelation of the sorts of interaction problems with which moral principles have to deal.⁷⁴ Suppose that individuals *J* and *K* each have a choice of two actions, *A* and *B*, and that *J*'s and *K*'s utilities are as follows:

		K	
		A	B
J	A	3,3	0,4
	B	4,0	2,2

In this case, *J* and *K* have a dominant strategy: to play *B*, and one has a prisoners' dilemma game. Each pursuing what is best for himself or herself results in a suboptimal outcome. Prisoners' dilemmas have proven to be not only a particularly salient way of representing problems of social cooperation, free-riding, and public goods provision, but they have also been a fruitful tool for formulating problems concerning moral cooperation. For example, Derek Parfit (1979) has argued that the underlying prisoners' dilemma structure arises not only in contrasting moral to self-interested action,

⁷⁴ For example, Hardin (1988) argues that an appreciation of the difficulties of strategic interaction throws a wrench into both simple applications and criticisms of utilitarianism. A sophisticated utilitarianism that recognizes these problems may be able to defend certain constellations of rights assignments as devices to simplify strategic interactions and to avoid suboptimal outcomes.

but also within the structure of moral thought itself. Each mother trying to save her own child from drowning in a beach accident, for example, may do less well than all would do if they cooperated in saving all the children. The collectively self-defeating character of moral commitments that are subject to this problem (like the paradoxes of humanitarianism discussed above, p. 677) is disturbing, and one might want to require that acceptable moral views not be collectively self-defeating. But many familiar moral maxims seem to fail in just this way.

Not all problems of social interaction have the structure of a prisoners' dilemma. There are also pure coordination problems (David Lewis 1969; Schelling 1978), such as determining which side of the road to drive on. These problems are easier in the sense that they may have self-enforcing solutions that can arise spontaneously. But they also have less bearing on moral problems, in which there are conflicts of interest to resolve. Conflicts of interest, however, are not always prisoners' dilemmas (Michael Taylor 1987, ch. 2; Hampton 1987; Hardin 1982). For example, suppose that *J*'s and *K*'s utilities are as follows:

		K	
		A	B
J	A	4,4	0,2
	B	2,0	1,1

For example, *J* and *K* might both want to do their part toward providing some collective good that requires both their efforts. One has here an "Assurance Problem" (Sen 1967; 1973b, pp. 96-99), in which the obstacles in the way of social cooperation are slighter than in a prisoners' dilemma and in which the prospects

for reconciling individual rationality and social optimality are much brighter. In fact, one strategy for coping with the social consequences of prisoners' dilemmas is to convert them into assurance problems by changing either incentives or personal motivations.

Suppose instead that the utilities look like this:

		K	
		A	B
J	A	3,3	2,4
	B	4,2	0,0

As in the assurance game *J* and *K* have no dominant strategy, but now they prefer to play *A* if and only if the other plays *B*. This game of "chicken" can arise easily when, for example, conflicts over possessions are costly. Both *J* and *K* might like the same fishing ground all to themselves. But sharing is better than ceding the fishing ground to the other, and fighting is worst of all. In such games, cooperation may be individually rational, particularly if, as Sugden (1986) has argued, the positions of the interactors are not fully symmetrical, so that conventions can arise directing those who occupy different positions to play different strategies. Property rights might be regarded as just such conventions, although they have of course been shaped by explicit legislation.

Indeed even a prisoners' dilemma structure of interactions need not prevent self-interested rational agents from cooperating if the interactions are iterated. Suppose that players in a prisoners' dilemma game are likely to play again. Defection is no longer a dominant strategy, as the ersatz cooperation of these entirely self-interested agents can elicit

like cooperation from others. Though the formal analysis of repeated games is complex (Michael Taylor 1987, ch. 3, 4), computer simulations have shown that simple strategies such as "tit-for-tat" (cooperate on the first move, then do whatever one's opponent did on the previous move do well against a wide range of alternatives (Robert Axelrod 1984). Indeed, provided that the probability of repeated play is high enough, a small group of individuals who play tit-for-tat can "invade" a population of players who defect on every move and, if "survival" depends on success, eventually drive those who always defect to extinction.⁷⁵ So cooperation can evolve among rational self-interested agents, and one might be tempted to regard much of morality as the conventional outcome of iterated games with various structures.

But one may question how useful game theory really is for the purposes of moral philosophy. Not only are there legitimate concerns about the assumption of common knowledge, but the use of game theory to address ethical questions also runs afoul of Sen's critique of welfarism—of the reliance on nothing but information about the utilities of outcomes. Game theory is not strictly welfarist, for non-preference information determines what the possible strategies are and what their utility consequences will be. But one might reasonably say that it is quasi-welfarist, for other ethically relevant factors—rights, capabilities, needs, and reasons—make no explicit appearance, and their role is heavily truncated (Roemer 1986a). For example, if the strategic problem is to distribute some set of commodities, then the solution will be the

⁷⁵ Tit-for-tat is in these circumstances an evolutionarily stable strategy. The biological terminology is not merely metaphor. Animal traits and behavior can be regarded as strategies, and game theory can be applied to the study of evolution. See Axelrod (1984, ch. 5).

same regardless of what the commodities are, if the feasible allocations and utility functions are the same. But people's judgments about just distributions depend on what the goods to be distributed are, and, as we shall see below the outcomes reached in experimental tests of bargaining theory depend on nonwelfarist information.⁷⁶

Ethical applications of game theory are also going to be precarious, because game theory itself is in turmoil. If controversy simmers in the theory of rational choice in circumstances of uncertainty, it boils in game theory. Not only are there the problems involved in incorporating imperfect knowledge and imperfect rationality, but there is disagreement about what solution concepts are appropriate, even in circumstances of perfect knowledge and rationality. The traditional solution concept for noncooperative game theory is the Nash equilibrium.⁷⁷ Most game theorists (for one exception, see Giacomo Bonnano 1991) would continue to hold that it is a reasonable necessary condition that solutions should be Nash equilibria. But it is not a sufficient condition: not all Nash equilibria are rational solutions (Reinhard Selten 1975). But there is no agreement on what stronger solution concepts to accept.

Furthermore, standard game theory has paradoxical implications. In a finitely iterated prisoners' dilemma game with a known number of iterations, the so-called backwards-induction argument apparently demonstrates that the uniquely rational strategy is defection on every

move. Real people do not play this way, and they wind up doing much better than if they played this purportedly uniquely rational strategy. The practical irrelevance of the backward induction argument may conceivably be explained by uncertainty about the number of plays parties will be engaged in. But experiments show that if individuals are permitted to form personal connections (which do not, however, involve any promises or contractual arrangements concerning their play), then they typically cooperate even in a one-shot prisoners' dilemma game (Dawes, van de Kragt, and Orbell 1990). The backwards induction argument is also itself puzzling, because it relies on apparently inconsistent counterfactual reasoning. Given common knowledge of perfect rationality, one is supposed to consider how an opponent would behave in a counterfactual situation that could not be reached by perfectly rational players with common knowledge of their perfect rationality (David Kreps et al. 1982; David Kreps and Robert Wilson 1982; Binmore 1987, 1988; Cristina Bicchieri 1988, 1990; Philip Pettit and Sugden 1989; Bonnano 1991). Most of game theory is currently too controversial to permit drawing ethical implications with any confidence.

Let us focus on one sort of game theory, which is subject to controversies of its own, but which has been applied to ethical issues: two-person cooperative bargaining theory. In John Nash's classic analysis, the two parties are supposed to have cardinal and interpersonally non-comparable utilities. The bargaining problem is to select a point from a closed convex set, S , of utility outcomes, containing a threat point, d , consisting of the utilities the two individuals receive if no bargain is reached. Nash proved that if the solution to the bargaining problem satisfies four explicit conditions (Pareto optimality, symmetry, indepen-

⁷⁶ Indeed Menachem Yaari and Maja Bar Hillel (1984) show that people have different ethical views about how to distribute avocados and grapefruits depending on whether they are told that the fruits are wanted for their taste or for their vitamins.

⁷⁷ A set of strategies is a Nash equilibrium if each strategy is a best reply to the others—that is, if no one can benefit by playing some other strategy given the strategies of others.

dence of irrelevant alternatives, and independence of positive linear utility transformations), then the solution maximizes the product of the utility gains from bargaining.⁷⁸ Although the Nash solution has many defenders, it has awkward consequences, for if the feasible set of utility outcomes expands in person A's favor, the result may be that person A will get less than before (Elster 1989a, p. 66). The Nash solution may also be morally objectionable in the extent to which it disadvantages the poor and risk averse.

A further problem is that, in empirical studies, the bargains individuals reach systematically diverge from the Nash solution, even when both individuals have perfect knowledge of their utility functions. In so-called binary lottery games, individuals bargain over the distribution of lottery tickets. Because the utility functions are not interpersonally comparable, one can stipulate that each individual's utility is zero when he or she has no tickets—that is, no probability of winning—and that the utility for each of winning the prize is 1. The utility of each player then equals the percentage of the lottery tickets each has. The Nash solution (this is the Kalai and Smorodinsky solution, too) is a fifty-fifty split of the lottery tickets. This is true regardless of whether the prizes the individuals can win are the same. When the prizes differ, the actual solutions to the bargaining games are bimodal with both the Nash solution and the solution that equalizes monetary returns as the two modes (Alvin Roth and Michael Malouf 1979). Roth, Malouf, and J. Keith Murnighan

(1981) argue persuasively that moral norms concerning just distributions affect the bargaining solutions. The person who stands to win the larger prize is seen as having a weaker moral claim on lottery tickets than the less advantaged person. So it seems that even purely positive bargaining theory must take cognizance of the moral commitments of the bargainers, and it is by no means obvious how this is to be accomplished formally (Gibbard 1990, esp. p. 262; Pettit 1990).

An alternative to the Nash solution developed by Kalai and Smorodinsky forms the core of Gauthier's (1986) ambitious contractarian effort to derive justice from rational choice. Kalai and Smorodinsky substitute a monotonicity requirement for Nash's independence of irrelevant alternatives, and their solution is called by Gauthier "minimax relative concession." The relative concession of each bargainer is $(u_{\max} - u)(u_{\max} - u_d)$, where u_{\max} is the largest feasible utility attainable by the individual, u is the utility if a given bargain were made, and u_d is the utility at the threat point—the utility that would result if no bargain were struck. Minimax relative concession minimizes the largest concession anybody has to make. If equal relative concession is Pareto optimal, it coincides with minimax relative concession. Gauthier argues that rational individuals who are aware of their own rationality and the rationality of others and who are aware of the utility consequences of their bargaining will accept the minimax relative concession solution (1986, ch. 5). In Gauthier's view this argument suffices to establish that such a solution is just, although Gauthier also maintains that this solution satisfies conditions that have been thought reasonable to impose on conceptions of justice, such as impartiality (1986, ch. 8). As we discussed on page 710 above, Gauthier argues further that it is individually rational to constrain

⁷⁸ Nash (1950). Luce and Raiffa (1957, pp. 124–37). A valuable recent study which emphasizes the normative uses of bargaining theory is William Thomson and Terje Lensberg (1989). In a striking recent development, Binmore (1990) shows how utilitarianism can be an evolutionarily stable solution to Nash's bargaining game with less than perfectly rational agents (see p. 710 and footnote 75 above).

one's individual maximization and to adhere to such bargains (1986, ch. 6).

Gauthier's argument has been strongly contested⁷⁹, and its difficulties highlight the complex interconnections between game theory and moral philosophy. For the characterization of individual rationality that Gauthier relies on is controversial, the relevance of solutions to a bargaining problem under conditions of complete information questionable, the particular solution concept dubious, and the limitation to quasi-welfarist information objectionable. This list of difficulties is not intended as a criticism of Gauthier's pathbreaking efforts or as an argument against employing game theory to tackle moral problems. Indeed it is only through Gauthier's efforts that these dimensions of a contractualist argument for principles of justice could be clearly seen. Even though game theory is in no condition at present to solve moral problems, it may some day do so, and even in its current state, it provides a rigorous setting in which to pose and ponder the problems.

5. Concluding Remarks

In the last twenty years, economists and moral philosophers have renewed a conversation that was interrupted during the heyday of positivist methodology in both disciplines. Although there remain considerable gaps between the modes of expression and habits of thought of moral philosophers and economists—as those who have read this far can no doubt confirm—we hope our presentation has also shown that **there is considerable room for productive interdisciplinary communication.** For those who aim to make scholarly contributions in this area, our

⁷⁹ See Kraus and Coleman (1987), Alan Nelson (1988), Ellen Paul et al. (1988), and also the more technical critiques of Roemer (1986a) and Sugden (1990).

review suggests a number of promising avenues of inquiry. These range from relatively technical investigations in areas such as game theory and social choice to conceptual and empirical questions about morally important notions such as equality and well-being. **Although we have not stressed them here, applications of moral reasoning to controversial areas of public policy, including for example affirmative action, environmental regulation, and social welfare measures, are also lively areas of inquiry.**

A better appreciation of the connections between moral theory and economics may also prove helpful to the many economists who will never do research on moral subjects directly. A clearer understanding of the moral perspectives that lie in the background of much of economic analysis can help economists to address the value-relevant aspects of their work. What Marshall said about economics—that it is “not a body of concrete truth, but an engine for the discovery of concrete truth” (1885, p. 159)—may be said about moral reasoning as well. An economics that is engaged actively and self-critically with the moral aspects of its subject matter cannot help but be more interesting, more illuminating, and ultimately more useful than one that tries not to be.

REFERENCES

- ACKERMAN, BRUCE. *Social justice in the liberal state*. New Haven: Yale U. Press, 1981.
- AKERLOF, GEORGE. “Loyalty Filters,” *Amer. Econ. Rev.*, Mar. 1983, 73(1), pp. 54–63.
- . “Gift Exchange and Efficiency-Wage Theory: Four Views,” *Amer. Econ. Rev.*, May 1984, 74(2), pp. 79–83.
- AKERLOF, GEORGE AND DICKENS, WILLIAM T. “The Economic Consequences of Cognitive Dissonance,” *Amer. Econ. Rev.*, June 1982, 72(3), pp. 307–19.
- AKERLOF, GEORGE AND YELLEN, JANET. “Introduction,” in *Efficiency wage models of the labor market*. Eds.: GEORGE AKERLOF AND JANET YELLEN. NY: Cambridge U. Press, 1986, pp. 1–21.
- ALEXANDER, SIDNEY S. “Social Evaluation through

- Notional Choice," *Quart. J. Econ.*, Nov. 1974, 88(4), pp. 597-624.
- ALLAIS, MAURICE. "The Foundations of a Positive Theory of Choice Involving Risk and a Criticism of the Postulates and Axioms of the American School," 1952; reprinted in *Expected utility hypotheses and the Allais paradox*. Eds.: MAURICE ALLAIS AND OLE HAGEN. Dordrecht: Reidel, 1979, pp. 27-145.
- ARNESON, RICHARD. "Equality and Equal Opportunity for Welfare," *Philosophical Studies*, May 1989, 56(1), pp. 77-93.
- . "Liberalism, Distributive Subjectivism, and Equal Opportunity for Welfare," *Philosophy and Public Affairs*, Spring 1990, 19(2), pp. 158-94.
- ARROW, KENNETH. *Social choice and individual values*. NY: Wiley, 1951; 2nd ed., 1963.
- . "Gifts and Exchanges," *Philosophy and Public Affairs*, Summer 1972, 1(4), pp. 343-62.
- . "Some Ordinalist-Utilitarian Notes on Rawls's Theory of Justice," *J. Philosophy*, May 1973, 70(9), pp. 245-63.
- . *The limits of organization*. 1st ed. NY: Norton, 1974.
- . "Extended Sympathy and the Possibility of Social Choice," *Philosophia*, June 1978, 7, pp. 223-37.
- D'ASPROMONT, CLAUDE AND GEVERS, LOUIS. "Equity and the Informational Basis of Collective Choice," *Rev. Econ. Stud.*, June 1977, 44(2), pp. 199-209.
- ATKINSON, ANTHONY B. "On the Measurement of Inequality," *J. Econ. Theory*, Sept. 1970, 2(3), pp. 244-63.
- ATKINSON, ANTHONY B. AND STIGLITZ, JOSEPH E. *Lectures on public economics*. NY: McGraw-Hill, 1980.
- AXELROD, ROBERT. *The evolution of cooperation*. NY: Basic Books, 1984.
- BAIER, KURT. *The moral point of view: A rational basis of ethics*. Ithaca: Cornell U. Press, 1958.
- BAKER, GEORGE P.; JENSEN, MICHAEL C. AND MURPHY, KEVIN J. "Compensation and Incentives: Practice vs. Theory," *J. Finance*, July 1988, 43(30), pp. 593-616.
- BANFIELD, EDWARD. *The moral basis of a backward society*. Glencoe, IL: Free Press, 1958.
- BARRY, BRIAN M. *Political argument*. London: Routledge, 1965.
- . *A treatise on social justice*. Vol. 1. *Theories of justice*. Berkeley, CA: U. of California Press, 1989.
- BASU, KAUSHIK. "The Right to Give Up Rights," *Economica*, Nov. 1984, 51(204), pp. 413-22.
- BAUMOL, WILLIAM. *Superfairness: Applications and theory*. Cambridge, MA: MIT Press, 1986.
- BECKER, GARY. *A treatise on the family*. Cambridge, MA: Harvard U. Press, 1981.
- BEITZ, CHARLES. "Review of Amartya Sen's *Resources, Values, and Development*," *Econ. Philos.*, Oct. 1986, 2(2), pp. 282-91.
- BENTHAM, JEREMY. *An introduction to the principles of morals and legislation*. Ed.: W. HARRISON. Oxford: Basil Blackwell, [1789] 1967.
- BERGSON, ABRAM. "A Reformulation of Certain Aspects of Welfare Economics," *Quart. J. Econ.*, Feb. 1938, 52(2), pp. 310-34.
- BERLIN, ISAIAH. "Two Concepts of Liberty," in *Four essays on liberty*. Oxford: Oxford U. Press, 1969, pp. 118-72.
- BICCHIERI, CRISTINA. "Self-Refuting Theories of Strategic Interaction: A Paradox of Common Knowledge," *Erkenntnis*, Feb. 1988, 29(1), pp. 69-85.
- . "Paradoxes of Rationality," *Midwest Studies in Philosophy*, 1990, 15, pp. 65-79.
- BINMORE, KEN. "Modeling Rational Players: Part I," *Econ. Philos.*, Oct. 1987, 3(2), pp. 179-214.
- . "Modeling Rational Players: Part II," *Econ. Philos.*, Apr. 1988, 4(1), pp. 9-56.
- . "Evolution and Utilitarianism," *Constitutional Political Economy*, Spring/Summer 1990, 1(2), pp. 1-26.
- BLACKORBY, CHARLES AND DONALDSON, DAVID. "Cash versus Kind, Self-selection, and Efficient Transfers," *Amer. Econ. Rev.*, Sept. 1988, 78(4), pp. 691-700.
- BLAU, JULIAN H. "Liberal Values and Independence," *Rev. Econ. Stud.*, July 1975, 42(3), pp. 395-401.
- BLINDER, ALAN S. AND CHOI, DON H. "A Shred of Evidence on Theories of Wage Stickiness," *Quart. J. Econ.*, Nov. 1990, 105(4), pp. 1003-15.
- BLOCK, WALTER. "A Free Market in Roads," in *MACHAN 1982*, pp. 164-83.
- BONNANO, GIACOMO. "The Logic of Rational Play in Games of Perfect Information," *Econ. Philos.*, Apr. 1991 7(1), pp. 37-66.
- BOULDING, KENNETH. *The economy of love and fear: A preface to grants economics*. Belmont, CA: Wadsworth, 1973.
- BOWLES, SAMUEL. "The Production Process in a Competitive Economy: Walrasian, Neo-Hobbesian and Marxian Models," *Amer. Econ. Rev.*, Mar. 1985, 75(1), pp. 16-36.
- BOWLES, SAMUEL AND GINTIS, HERBERT. "The Democratic Firm: An Agency-Theoretic Evaluation." Amherst: Dept. of Economics, U. of Massachusetts, May 1990.
- BRANDT, RICHARD B. *A theory of the right and the good*. Oxford: Oxford U. Press, 1979.
- BRAVERMAN, HARRY. *Labor and monopoly capital*. NY: Monthly Review Press, 1974.
- BRAYBROOKE, DAVID. *Meeting needs*. Princeton: Princeton U. Press, 1987.
- BRENNAN, GEOFFREY AND BUCHANAN, JAMES. *The reason of rules: Constitutional political economy*. NY: Cambridge U. Press, 1985.
- BROOME, JOHN. "Should Social Preferences Be Consistent?" *Econ. Philos.*, Apr. 1989, 5(1), pp. 7-17.
- . "Utility," *Econ. Philos.*, Apr. 1991a, 7(1), pp. 1-12.
- . *Weighing goods*. Oxford: Basil Blackwell, 1991b.
- BUCHANAN, ALLEN E. *Ethics, efficiency, and the market*. Totowa, NJ: Rowman & Allanheld, 1985.
- . "Assessing the Communitarian Critique of Liberalism," *Ethics*, July 1989, 99(4), pp. 852-82.
- BUCHANAN, JAMES M. *The limits of liberty: Between*

- anarchy and leviathan*. Chicago: U. of Chicago Press, 1975.
- BUCHANAN, JAMES M. AND TULLOCK, GORDON. *The calculus of consent*. Ann Arbor: U. of Michigan Press, 1962.
- CALABRESI, GUIDO AND MELAMED, DOUGLAS. "Property Rules, Liability Rules, and Inalienability: One View of the Cathedral," *Harvard Law Rev.*, Apr. 1972, 85(6), pp. 1089–1128.
- CARENS, JOSEPH. "Compensatory Justice and Social Institutions," *Econ. Philos.*, Apr. 1985, 1(1), pp. 39–67.
- COASE, RONALD. "The Problem of Social Cost," *J. Law Econ.*, Oct. 1960, 3(1), pp. 1–44.
- COHEN, GERALD ALLEN. "Functional Explanation, Consequence Explanation, and Marxism," *Inquiry*, Mar. 1982, 25(1), pp. 27–56.
- COHEN, JOSHUA. "An Epistemic Conception of Democracy," *Ethics*, Oct. 1986, 97(1), pp. 26–38.
- . "Deliberation and Democratic Legitimacy," in *The good polity: Normative analysis of the state*. Eds.: ALAN HAMLIN AND PHILIP PETTIT. NY: Basil Blackwell, 1989, pp. 17–34.
- COLEMAN, JULES AND FERREJOHN, JOHN. "Democracy and Social Choice," *Ethics*, Oct. 1986, 97(1), pp. 6–25.
- COLEMAN, JULES L. AND KRAUS, JODY. "Rethinking the Theory of Legal Rights," *Yale Law Rev.*, June 1986, 95(7), pp. 1335–71; reprinted in *Markets, morals and the law*. Cambridge: Cambridge U. Press, 1988, pp. 28–63.
- COLLARD, DAVID. *Altruism and economy: A study in non-selfish economics*. NY: Oxford U. Press, 1978.
- COOTER, ROBERT. "Rawls's Lexical Orderings Are Good Economics," *Econ. Philos.*, Apr. 1989, 5(1), pp. 47–54.
- COOTER, ROBERT AND RAPPOPORT, PETER. "Were the Ordinalists Wrong About Welfare Economics?" *J. Econ. Lit.*, June 1984, 22(2), pp. 507–30.
- DANIELS, NORMAN. "Wide Reflective Equilibrium and Theory Acceptance in Ethics," *J. Philosophy*, May 1979, 76(5), pp. 256–82.
- DASGUPTA, PARTHA. *The control of resources*. Cambridge: Harvard U. Press, 1982.
- DASGUPTA, PARTHA; HAMMOND, PETER AND MASKIN, ERIC. "The Implementation of Social Choice Rules: Some General Results on Incentive Compatibility," *Rev. Econ. Stud.*, Apr. 1979, 46(143), pp. 185–216.
- DAWES, ROBYN; VAN DE KRACHT, ALPHONS AND ORBELL, JOHN. "Cooperation for the Benefit of Us—Not Me, or My Conscience," in *Beyond self-interest*. Ed.: JANE J. MANSBRIDGE. Chicago: U. of Chicago Press, 1990, pp. 97–110.
- DEBBREU, GERARD. *Theory of value*. NY: Wiley, 1959.
- DORE, RONALD. *Taking Japan seriously: A Confucian perspective on leading economic issues*. London: The Athlone Press, 1987.
- DUNLOP, JOHN T. "The Task of Contemporary Wage Theory," in *The theory of wage determination*. Ed.: JOHN T. DUNLOP. NY: St. Martin's Press, 1957, pp. 3–27.
- DWORKIN, GERALD. "Paternalism," in *Morality and the law*. Ed.: RICHARD WASSERSTROM. Belmont, CA: Wadsworth Pub. Co., 1971, pp. 107–36.
- . *The theory and practice of autonomy*. Cambridge: Cambridge U. Press, 1988.
- DWORKIN, GERALD; BERMANT, GORDON AND BROWN, PETER G., eds. *Markets and morals*. Washington: Hemisphere Pub. Corp.; NY: distributed solely by Halsted Press, 1977.
- DWORKIN, RONALD. *Taking rights seriously*. Cambridge: Harvard U. Press, 1977.
- . "What is Equality? Part 1: Equality of Welfare," *Philosophy and Public Affairs*, Spring 1981a, 10(3), pp. 185–246.
- . "What is Equality? Part 2: Equality of Resources," *Philosophy and Public Affairs*, Summer 1981b, 10(4), pp. 283–345.
- . "Foundations of Liberal Equality," in *The Tanner lectures on human values*, XI. Ed.: GRETHE B. PETERSON. Salt Lake City: U. of Utah Press, 1990, pp. 1–119.
- EDWARDS, RICHARD. *Contested terrain: The transformation of the workplace in the twentieth century*. NY: Basic Books, 1979.
- ELLSBERG, DANIEL. "Risk, Ambiguity, and the Savage Axioms," *Quart. J. Econ.*, Nov. 1961, 75(4), pp. 643–99.
- ELSTER, JON. *Ulysses and the sirens: Studies in rationality and irrationality*. NY: Cambridge U. Press, 1979.
- . *Sour grapes: Studies in the subversion of rationality*. Cambridge: Cambridge U. Press, 1983.
- . "The Market and the Forum: Three Varieties of Political Theory," in ELSTER AND HYLLAND 1986, pp. 103–32.
- . *The cement of society: A study of social order*. Cambridge: Cambridge U. Press, 1989a.
- . "Social Norms and Economic Theory," *J. Econ. Perspectives*, Fall 1989b, 3(4), pp. 99–117.
- ELSTER, JON AND HYLLAND, AANUND, eds. *Foundations of social choice theory*. NY: Cambridge U. Press, 1986.
- ELSTER, JON AND ROEMER, JOHN, eds. *Interpersonal comparisons of well-being*. Cambridge: Cambridge U. Press, 1991.
- ESTLUND, DAVID. "Democracy Without Preference," *Philosophical Rev.*, July 1990, 99(3), pp. 397–423.
- ETZIONI, AMITAI. "The Case for a Multiple-Utility Conception," *Econ. Philos.*, Oct. 1986, 2(2), pp. 159–83.
- . *The moral dimension. Toward a new economics*. NY: Macmillan, 1988.
- FEINBERG, JOEL. *Social philosophy*. Englewood Cliffs, NJ: Prentice-Hall, 1973.
- . *Harm to self*. Oxford: Oxford U. Press, 1986.
- FLEMING, JOHN MARCUS. "A Cardinal Concept of Welfare," *Quart. J. Econ.*, Aug. 1952, 66(3), pp. 366–84.
- FOLEY, DUNCAN. "Resource Allocation and the Public Sector," *Yale Econ. Essays*, Spring 1967, 7, pp. 45–98.
- FOX, ALAN. *Beyond contract: Work, power and trust relations*. London: Faber, 1974.
- FRANK, ROBERT. *Passions within reason: The strategic*

- role of the emotions. NY: W. W. Norton, 1988.
- FRANKFURT, HARRY. "Freedom of Will and the Concept of a Person," *J. Philosophy*, Jan. 1971, 68(1), pp. 5-20.
- _____. "Equality as a Moral Ideal," *Ethics*, Oct. 1987, 98(1), pp. 21-43.
- FRIEDMAN, MILTON. *Capitalism and freedom*. Chicago: U. of Chicago Press, 1962.
- GAERTNER, WULF; PATTANAİK, PRASANTA AND SUZUMURA, KOTARO. "Individual Rights Revisited," *Economica*, forthcoming.
- GAMBETTA, DIEGO, ed. *Trust: Making and breaking cooperative relations*. Oxford: Basil Blackwell, 1988.
- GARDENFORS, PETER. "Rights, Games and Social Choice," *Nous*, Sept. 1981, 15(3), pp. 341-56.
- GAUTHIER, DAVID. "On the Refutation of Utilitarianism," in MILLER AND WILLIAMS 1982.
- _____. *Morals by agreement*. Oxford: Clarendon Press, 1986.
- GIBBARD, ALLAN. "A Pareto-Consistent Libertarian Claim," *J. Econ. Theory*, Apr. 1974, 7(4), pp. 388-410.
- _____. "Interpersonal Comparisons: Preference, Good, and the Intrinsic Reward of a Life," in ELSTER AND HYLLEND 1986, pp. 165-94.
- _____. *Wise choices, apt feelings: A theory of normative judgment*. Cambridge, MA: Harvard U. Press, 1990.
- GILBERT, MARGARET. "Walking Together: A Paradigmatic Social Phenomenon," in *Midwest Studies in Philosophy*, Vol. 15. *The Philosophy of the Human Sciences*. Eds.: PETER FRENCH, THEODORE Uehling, and HOWARD WETTSTEIN. Notre Dame, IN: U. of Notre Dame Press, 1990, pp. 1-14.
- GRAAFF, JAN DE VAN. *Theoretical welfare economics*. Cambridge: Cambridge U. Press, 1957.
- GRIFFIN, JAMES. *Well-being: Its meaning, measurement and moral importance*. Oxford: Clarendon Press, 1986.
- GUL, FAROUK A. "A Theory of Disappointment Aversion," *Econometrica*, May 1991, 59(3), pp. 667-86.
- HAMLIN, ALAN. *Ethics, economics, and the state*. NY: St. Martin's Press, 1986.
- HAMMOND, PETER J. "Changing Tastes and Coherent Dynamic Choice," *Rev. Econ. Stud.*, July 1976a, 43(1), pp. 159-73.
- _____. "Equity, Arrow's Conditions, and Rawls' Difference Principle," *Econometrica*, July 1976b, 44(4), pp. 793-804.
- _____. "Ex-Post Optimality as a Dynamically Consistent Objective for Collective Choice under Uncertainty," in *Social choice and welfare*. Eds.: PRASANTA PATTANAİK AND MAURICE SALLES. Amsterdam: North-Holland, 1983, pp. 175-205.
- _____. "Consequentialist Foundations for Expected Utility," *Theory and Decision*, Jan. 1988, 25(1), pp. 25-78.
- HAMPTON, JEAN. *Hobbes and the social contract tradition*. Cambridge: Cambridge U. Press, 1986.
- _____. "Free-Rider Problems in the Production of Collective Goods," *Econ. Philos.*, Oct. 1987, 3(2), pp. 245-73.
- HANDS, DOUGLAS WADE. "Ad Hocness in Economics and the Popperian Tradition," in *The Popperian legacy in economics*. Ed.: NEIL DE MARCHI. Cambridge: Cambridge U. Press, 1988, pp. 121-37.
- HARBERGER, ARNOLD C. "On the Use of Distributional Weights in Social Cost-Benefit Analysis," *J. Polit. Econ.*, Apr. 1978, 86(2), pp. S87-120.
- HARDIN, RUSSELL. *Collective action*. Baltimore: Johns Hopkins U. Press, 1982.
- _____. *Morality within the limits of reason*. Chicago: U. of Chicago Press, 1988.
- HARE, RICHARD M. *Moral thinking: Its levels, method, and point*. Oxford: Oxford U. Press, 1981.
- HARROD, ROY. "Utilitarianism Revised," *Mind*, Apr. 1936, 45(178), pp. 137-56.
- HARSANYI, JOHN. "Cardinal Welfare, Individualistic Ethics and Interpersonal Comparisons of Utility," *J. Polit. Econ.*, Aug. 1955, 63(4), pp. 309-21.
- _____. "Morality and the Theory of Rational Behavior," *Social Research*, Winter, 1977a, 44(4), pp. 623-56; reprinted in SEN AND WILLIAMS 1982, pp. 39-62.
- _____. *Rational behavior and bargaining equilibrium in games and social situations*. Cambridge: Cambridge U. Press, 1977b.
- HART, HERBERT L. A. "Are There any Natural Rights?" *Philosophical Review*, Apr. 1955, 64(2), pp. 175-91; reprinted in WALDRON 1984, pp. 77-90.
- HASLETT, D. W. "What is Utility?" *Econ. Philos.*, Apr. 1990, 6(1), pp. 65-94.
- HAUSMAN, DANIEL. *The inexact and separate science of economics*. Cambridge: Cambridge U. Press, 1992.
- HAYEK, FRIEDRICH VON. *The constitution of liberty*. Chicago: U. of Chicago Press, 1960.
- _____. "The Moral Element in Free Enterprise," in *Studies in philosophy, politics and economics*. Chicago: U. of Chicago Press, 1967, pp. 229-36.
- _____. *The mirage of social justice*. Chicago: U. of Chicago Press, 1976.
- _____. *The fatal conceit: The errors of socialism*. Vol. 1. *Collected works of F. A. Hayek*. London: Routledge, 1988.
- HERSTEIN, ISRAEL AND MILNOR, JOHN W. "An Axiomatic Approach to Measurable Utility," *Econometrica*, Apr. 1953, 21(2), pp. 291-97.
- HICKS, JOHN. "The Foundations of Welfare Economics," *Econ. J.*, Dec. 1939, 49(196), pp. 696-712.
- _____. "The Valuation of Social Income," *Economica N.S.*, May 1940, 7, pp. 105-24.
- HIRSCH, FRED. *Social limits to growth*. Cambridge: Harvard U. Press, 1976.
- HIRSCHMAN, ALBERT O. "Against Parsimony: Three Easy Ways of Complicating Some Categories of Economic Discourse," *Econ. Philos.*, Apr. 1985, 1(1), pp. 7-21.
- HOBBS, THOMAS. *Leviathan, or the matter, form and power of a commonwealth ecclesiastical and civil*. NY: Collier Books, [1651] 1962.
- HOCHMAN, HAROLD M. AND ROGERS, JAMES D. "Pareto Optimal Redistribution," *Amer. Econ. Rev.*, Sept. 1969, 59(4), pp. 542-57.

- HODGSON, GEOFFREY M. "Hayek's Theory of Cultural Evolution: An Evaluation in the Light of Vanberg's Critique," *Econ. Philos.*, Apr. 1991, 7(1), pp. 67-82.
- HOHFELD, WESLEY. *Fundamental legal conceptions*. New Haven: Yale U. Press, 1923.
- HURLEY, SUSAN L. *Natural reasons: Personality and polity*. Oxford: Oxford U. Press, 1989.
- JENSEN, NIELS. "An Introduction to Bernoullian Utility Theory: I. Utility Functions," *Swedish J. Econ.*, Sept. 1967, 69, pp. 163-83.
- KAGAN, SHELLY. *The limits of morality*. Oxford: Oxford U. Press, 1989.
- KALDOR, NICHOLAS. "Welfare Propositions of Economics and Interpersonal Comparisons of Utility," *Econ. J.*, Sept. 1939, 49(195), pp. 549-52.
- KANT, IMMANUEL. *Groundwork of the metaphysics of morals*. Trans.: H. PATON. NY: Harper & Row, [1785] 1948.
- KAVKA, GREGORY S. *Hobbesian moral and political theory*. Princeton: Princeton U. Press, 1986.
- KELMAN, STEVEN. "A Case for In-Kind Transfers," *Econ. Philos.*, Apr. 1986, 2(1), pp. 55-73.
- KNIGHT, FRANK H. *Risk, uncertainty and profit*. Chicago: U. of Chicago Press, [1921] 1971.
- . *The ethics of competition, and other essays*. NY and London: Harper & Brothers, 1935.
- KOLM, SERGE-CHRISTOPHE. *Justice et équité*. Paris: Editions du Centre National de la Recherche Scientifique, 1972.
- KRAUS, JODY AND COLEMAN, JULES. "Morality and the Theory of Rational Choice," *Ethics*, July 1987, 97(4), pp. 715-49.
- KREPS, DAVID M. ET AL. "Rational Cooperation in the Finitely Repeated Prisoners' Dilemma," *J. Econ. Theory*, Aug. 1982, 27(2), pp. 245-52.
- KREPS, DAVID AND WILSON, ROBERT. "Sequential Equilibria," *Econometrica*, July 1982, 50(4), pp. 863-94.
- LEIBENSTEIN, HARVEY. *Beyond economic man: A new foundation for economics*. Cambridge, MA: Harvard U. Press, 1976.
- . "A Branch of Economics is Missing: Micro-Micro Theory," *J. Econ. Lit.*, June 1979, 17(2), pp. 477-502.
- LEARNER, ABBA. "Consumption-Loan Interest and Money," *J. Polit. Econ.*, Oct. 1959, 67(5), pp. 512-25.
- LEVI, ISAAC. *The enterprise of knowledge*. Cambridge, MA: MIT Press, 1980.
- . *Hard choices: Decision-making under unresolved conflict*. Cambridge: Cambridge U. Press, 1986a.
- . "The Paradoxes of Allais and Ellsberg," *Econ. Philos.*, Apr. 1986b, 2(1), pp. 23-53.
- . "Pareto Unanimity and Consensus," *J. Philosophy*, Sept. 1990, 89(9), pp. 481-92.
- LEWIS, DAVID. *Convention. A philosophical study*. Cambridge: Harvard U. Press, 1969.
- LIND, ROBERT C. ET AL. *Discounting for time and risk in energy policy*. Washington, DC: Resources for the Future; distributed by Johns Hopkins U. Press, Baltimore, 1982.
- LIPSEY, RICHARD AND LANCASTER, KELVIN. "The General Theory of Second Best," *Rev. Econ. Stud.*, Oct. 1956, 24(1), pp. 11-32.
- LITTLE, IAN. *A critique of welfare economics*. 2nd ed. Oxford: Oxford U. Press, 1957.
- LOMASKY, LOREN. *Persons, rights and the moral community*. NY: Oxford U. Press, 1987.
- LOOMES, GRAHAM AND SUGDEN, ROBERT. "Regret Theory: An Alternative Theory of Rational Choice under Uncertainty," *Econ. J.*, Dec. 1982, 92(368), pp. 805-24.
- LUCE, ROBERT DUNCAN AND RAIFFA, HOWARD. *Games and decisions*. NY: Wiley, 1957.
- LYONS, DAVID. "Human Rights and the General Welfare," *Philosophy and Public Affairs*, Winter 1977, 6(1), pp. 113-29.
- MACHAN, TIBOR R., ed. *The libertarian reader*. Totowa, NJ: Rowman & Littlefield, 1982.
- MACHINA, MARK. "Choice under Uncertainty: Problems Solved and Unsolved," *J. Econ. Perspectives*, Summer 1987, 1(1), pp. 121-54.
- . "Dynamic Consistency and Non-Expected Utility Models of Choice Under Uncertainty," *J. Econ. Lit.*, Dec. 1989, 27(4), pp. 1622-68.
- MACHLUP, FRITZ. "Positive and Normative Economics," in *Economic means and social ends: Essays in political economics*. Ed.: ROBERT HEILBRONER. Englewood Cliffs, NJ: Prentice-Hall, 1969, pp. 99-129.
- MACINTYRE, ALASTAIR C. *After virtue: A study in moral theory*. London: Duckworth, 1981.
- MACKAY, ALFRED. "Extended Sympathy and Interpersonal Utility Comparisons," *J. Philosophy*, June 1986, 83(6), pp. 305-22.
- MACKIE, JOHN. *Ethics: Inventing right and wrong*. Harmondsworth: Penguin, 1977.
- MACLEAN, DOUGLAS AND BROWN, PETER G., eds. *Energy and the future*. Totowa, NJ: Rowman & Littlefield, 1983.
- MANSBRIDGE, JANE J. "Expanding the Range of Formal Modeling," in *Beyond self-interest*. Ed.: JANE MANSBRIDGE. Chicago: U. of Chicago Press, 1990, pp. 254-63.
- MARGOLIS, HOWARD. *Selfishness, altruism, and rationality*. Cambridge: Cambridge U. Press, 1982.
- MARSHALL, ALFRED. "The Present Position of Economics," 1885; reprinted in *Memorials of Alfred Marshall*. Ed.: A. C. PIGOU. London: Macmillan, 1925, pp. 152-74.
- . "A Fair Rate of Wages," 1887; reprinted in *Memorials of Alfred Marshall*. Ed.: A. C. PIGOU. London: Macmillan, 1925, pp. 212-26.
- MARWELL, GERALD AND AMES, RUTH. "Economists Free Ride. Does Anyone Else? Experiments on the Provision of Public Goods. IV," *J. Pub. Econ.*, June 1981, 15(3), pp. 295-310.
- MASKIN, ERIC. "The Theory of Implementation in Nash Equilibrium," in *Social goals and social organization: Essays in memory of Elisha Pazner*. Eds.: LEONID HURWICZ, DAVID SCHMEIDLER, AND HUGO SONNENSCHNEIN. Cambridge: Cambridge U. Press, 1985, pp. 173-204.
- MCCLENNEN, EDWARD. *Rationality and dynamic choice: Foundational explorations*. Cambridge: Cambridge U. Press, 1990.

- McDOWELL, JOHN. "Are Moral Requirements Hypothetical Imperatives?" *Proceedings of the Aristotelian society supplementary volume*. 1978, pp. 13-29.
- McKEAN, ROLAND N. "Economics of Trust, Altruism, and Corporate Responsibility," in *Altruism, morality and economic theory*. Ed.: EDMUND S. PHELPS. NY: Russell Sage Foundation, 1975, pp. 29-44.
- McPHERSON, MICHAEL. "Mill's Moral Theory and the Problem of Preference Change," *Ethics*, Jan. 1982, 92(1), pp. 252-73.
- _____. "Efficiency and Liberty in the Productive Enterprise: Recent Work in the Economics of Work Organization," *Philosophy and Public Affairs*, Fall 1983a, 12(4), pp. 354-68.
- _____. "Want Formation, Morality, and Some 'Interpretive' Aspects of Economic Inquiry," in *Social science as moral inquiry*. Eds.: NORMA HAAN ET AL. NY: Columbia U. Press, 1983b, pp. 96-124.
- _____. "Limits on Self-seeking: The Role of Morality in Economic Life," in *Neoclassical political economy*. Ed.: DAVID COLANDER. Boston: Ballinger Press, 1984, pp. 71-85.
- MECKLING, WILLIAM. "An Exact Consumption-Loan Model of Interest: A Comment," *J. Polit. Econ.*, Feb. 1960, 68(1), pp. 72-76.
- MILL, JOHN STUART. *On liberty*. 1859. Ed.: CURRIN V. SHIELDS. NY: Macmillan, 1985.
- _____. *Principles of political economy*. 7th ed. Ed.: WILLIAM J. ASHLEY [1871], 1909; reprinted NY: A. M. Kelley, 1976.
- MILLER, HARLAN B. AND WILLIAMS, WILLIAM H. eds. *The limits of utilitarianism*. Minneapolis: U. of Minnesota Press, 1982.
- MIRRELES, JAMES A. "Optimal Tax Theory: A Synthesis," *J. Pub. Econ.*, Nov. 1976, 6(4), pp. 327-58.
- MORISHIMA, MICHIO. *Why has Japan succeeded?* NY: Cambridge U. Press, 1982.
- MOORE, GEORGE EDWARD. *Principia ethica*. Cambridge: Cambridge U. Press, 1902; reprinted 1959, 1984.
- MUELLER, DENNIS. "Intergenerational Justice and the Social Discount Rate," *Theory and Decision*, Oct. 1974, 5(3), pp. 263-73.
- MUSGRAVE, RICHARD. "Maximin, Uncertainty, and the Leisure Trade-Off," *Quart. J. Econ.*, Nov. 1974, 88(4), pp. 625-32.
- NAGEL, THOMAS. *The possibility of altruism*. Oxford: Clarendon Press, 1970.
- _____. "The Limits of Objectivity," in *The Tanner lectures on human values*. Vol. I. Ed.: STERLING M. McMURRIN. Salt Lake City: U. of Utah Press, 1980, pp. 75-139.
- _____. "Libertarianism Without Foundations," in *Reading Nozick: Essays on anarchy, state, and utopia*. Ed.: JEFFREY PAUL. Oxford: Basil Blackwell, 1981, pp. 191-205.
- NARVESON, JAN. *The libertarian idea*. Philadelphia: Temple U. Press, 1988.
- NASH, JOHN. "The Bargaining Problem," *Econometrica*, Jan. 1950, 18(1), pp. 155-62.
- NELSON, ALAN. "Economic Rationality and Morality," *Philosophy and Public Affairs*, Spring 1988, 17(2), pp. 149-66.
- VON NEUMANN, JOHN AND MORGENSTERN, OSKAR. *Theory of games and economic behavior*. 2nd ed. Princeton: Princeton U. Press, 1947.
- NOZICK, ROBERT. *Anarchy, state, and utopia*. NY: Basic Books, 1974.
- NUTZINGER, HANS G. "The Firm as a Social Institution: The Failure of a Contractarian Viewpoint." Working Paper No. 52, Alfred Weber Institute, U. of Heidelberg, 1976.
- OKUN, ARTHUR. *Equality and efficiency: The big tradeoff*. Washington, DC: Brookings Institution, 1975.
- PARFIT, DEREK. "Prudence, Morality and the Prisoners' Dilemma," *Proceedings of the British Academy*, 1979.
- _____. *Reasons and persons*. Oxford: Oxford U. Press, 1984.
- VAN PARIJS, PHILIPPE. "On the Ethical Foundations of Basic Income." Working Paper #CMID 32, Institut Supérieur de Philosophie, Université Catholique de Louvain, 1989.
- _____. "The Second Marriage of Justice and Efficiency," *J. Social Policy*, 1990, 19, pp. 1-25.
- _____. "Why Surfers Should Be Fed: The Liberal Case for an Unconditional Basic Income," *Philosophy and Public Affairs*, Spring 1991, 20(2), pp. 101-31.
- PAUL, ELLEN FRANKEL ET AL., eds. *The new social contract: Essays on Gauthier*. Oxford: Blackwell, 1988.
- PELEG, BEZALEL. *Game theoretic analysis of voting in committees*. Cambridge: Cambridge U. Press, 1984.
- PETTIT, PHILIP. "Virtus Normativa: Rational Choice Perspectives," *Ethics*, July 1990, 100(4), pp. 725-55.
- PETTIT, PHILIP AND SUGDEN, ROBERT. "The Backward Induction Paradox," *J. Philosophy*, Apr. 1989, 86(4), pp. 169-82.
- PHELPS, EDMUND S. "Taxation of Wage Income for Economic Justice," *Quart. J. Econ.*, Aug. 1973, 87(3), pp. 331-54.
- POLYANI, KARL. *The great transformation*. New York and Toronto: Farrar & Rinehart, 1944.
- POSNER, RICHARD. "Utilitarianism, Economics, and Legal Theory," *J. Legal Stud.*, Jan. 1979, 8(1), pp. 103-40.
- PUTNAM, HILARY. *Reason, truth and history*. Cambridge: Cambridge U. Press, 1981.
- PUTTERMAN, LOUIS. "On Some Recent Explanations of Why Capital Hires Labor," *Econ. Inquiry*, Apr. 1984, 22(2), pp. 171-87.
- RAILTON, PETER. "Alienation, Consequentialism, and the Demands of Morality," *Philosophy and Public Affairs*, Spring 1984, 13(2), pp. 134-71.
- RAMSEY, FRANK. "Truth and Probability," in *The foundations of mathematics and other logical essays*. Ed.: R. B. BRAITHWAITE. London: Routledge & Kegan Paul, 1926, pp. 156-98.
- RAND, AYN. *The virtue of selfishness*. NY: Signet Books, 1964.
- RAWLS, JOHN. "Outline of a Decision Procedure for

- Ethics," *Philosophical Rev.*, Apr. 1951, 60(2), pp. 177-97.
- _____. "Two Concepts of Rules," *Philosophical Rev.*, Apr. 1955, 64(1), pp. 3-33.
- _____. *A theory of justice*. Cambridge, MA: Harvard U. Press, 1971.
- _____. "Social Unity and Primary Goods," in SEN AND WILLIAMS 1982, pp. 159-86.
- RAZ, JOSEPH. *The morality of freedom*. Oxford: Oxford U. Press, 1986.
- REDER, MELVIN. "The Place of Ethics in the Theory of Production," in *Economics and human welfare: Essays in honor of Tibor Scitovsky*. Ed.: MICHAEL BOSKIN. NY: Academic Press, 1979, pp. 133-46.
- REGAN, DONALD. *Utilitarianism and cooperation*. Oxford: Clarendon Press, 1980.
- _____. "Against Evaluator Relativity: A Response to Sen," *Philosophy and Public Affairs*, Spring 1983, 12(2), pp. 93-112.
- REICH, MICHAEL AND DEVINE, JAMES. "The Microeconomics of Conflict and Hierarchy in Capitalist Production," *Rev. Radical Polit. Econ.*, Winter 1981, 12(1), pp. 27-45.
- RILEY, JONATHAN. "Rights to Liberty in Purely Private Matters, Part I," *Econ. Philos.*, Oct. 1989, 5(2) pp. 121-66.
- _____. "Rights to Liberty in Purely Private Matters, Part II," *Econ. Philos.*, Apr. 1990, 6(1), pp. 27-64.
- ROBBINS, LIONEL. *An essay on the nature and significance of economic science*. 2nd ed. London: Macmillan, [1932] 1935.
- ROBINSON, JOAN. *The economics of imperfect competition*. London: Macmillan, 1933.
- ROEMER, JOHN. *A general theory of exploitation and class*. Cambridge, MA: Harvard U. Press, 1982.
- _____. "Equality of Talent," *Econ. Philos.*, Oct. 1985a, 1(2), pp. 151-88.
- _____. "Should Marxists Be Interested in Exploitation?" *Philosophy and Public Affairs*, Winter 1985b, 14(1), pp. 30-65.
- _____. "The Mismatch of Bargaining Theory and Distributive Justice," *Ethics*, Oct. 1986a, 97(1), pp. 88-110.
- _____. "Equality of Resources Implies Equality of Welfare," *Quart. J. Econ.*, Nov. 1986b, 101(4), pp. 751-84.
- _____. "Egalitarianism, Responsibility, and Information," *Econ. Philos.*, Oct. 1987, 3(2), pp. 215-44.
- _____. *Free to lose*. Cambridge: Harvard U. Press, 1988.
- _____. "What Is Exploitation? Reply to Jeffrey Reiman," *Philosophy and Public Affairs*, Winter 1989, 18(1), pp. 90-97.
- RORTY, RICHARD. *The consequences of pragmatism*. Minneapolis: U. of Minnesota Press, 1982.
- ROTH, ALVIN AND MALOUF, MICHAEL. "Game Theoretical Models and the Role of Information in Bargaining," *Psych. Rev.*, Nov. 1979, 86(6), pp. 574-94.
- ROTH, ALVIN; MALOUF, MICHAEL AND MURNIGHAN, J. KEITH. "Sociological versus Strategic Factors in Bargaining," *J. Econ. Behav. Organ.*, June 1981, 2(2), pp. 153-77.
- SAGOFF, MARK. "Values and Preferences," *Ethics*, Jan. 1986, 96(2), pp. 301-16.
- SAMUELSON, PAUL. *Foundations of economic analysis*. Cambridge, MA: Harvard U. Press, 1947.
- _____. "Evaluation of Real National Income," *Oxford Econ. Pap.*, N.S., Jan. 1950, 2(1), pp. 1-29.
- _____. "An Exact Consumption-Loan Model of Interest with or without the Social Contrivance of Money," *J. Polit. Econ.*, Dec. 1958, 66(6), pp. 467-82.
- _____. "Reaffirming the Existence of 'Reasonable' Bergson-Samuelson Social Welfare Functions," *Economica N.S.*, Feb. 1977, 44(173), pp. 81-88.
- SANDEL, MICHAEL. *Liberalism and the limits of justice*. Cambridge: Cambridge U. Press, 1982.
- SAVAGE, LEONARD. *The foundations of statistics*. NY: Dover, 1972.
- SCANLON, THOMAS. "Preference and Urgency," *J. Philosophy*, Nov. 1975, 75(22), pp. 655-70.
- _____. "Contractualism and Utilitarianism," in SEN AND WILLIAMS 1982, pp. 103-28.
- _____. "Equality of Resources and Equality of Welfare: A Forced Marriage?" *Ethics*, Oct. 1986, 97(1), pp. 111-18.
- SCHEFFLER, SAMUEL. *The rejection of consequentialism: A philosophical investigation of the considerations underlying rival moral conceptions*. Oxford: Clarendon Press, 1982.
- _____, ed. *Consequentialism and its critics*. Oxford: Oxford U. Press, 1988.
- SCHELLING, THOMAS. *Micromotives and macrobehavior*. NY: Norton, 1978.
- _____. *Choice and consequence*. Cambridge: Harvard U. Press, 1984.
- SCHICK, FREDERIC. "D Bookies and Money Pumps," *J. Philosophy*, Feb. 1986, 83(2), pp. 112-19.
- SCHMEIDLER, DAVID AND PAZNER, ELISHA A. "A Difficulty in the Concept of Fairness," *Rev. Econ. Stud.*, July 1974, 41(3), pp. 441-43.
- SCHOTTER, ANDREW. *The economic theory of social institutions*. Cambridge: Cambridge U. Press, 1981.
- SCHWARTZ, THOMAS. "What Welfare Is Not," in MILLER AND WILLIAMS 1982, pp. 195-206.
- SEIDENFELD, TEDDY; KADANE, JOSEPH AND SCHERVISH, MARK. "On the Shared Preferences of Two Bayesian Decision Makers," *J. Philosophy*, May 1989, 86(5), pp. 225-44.
- SELTEN, REINHART. "Re-examination of the Perfectness Concept for Equilibrium in Extensive Games," *Int. J. Game Theory*, 1975, 4, pp. 22-25.
- SEN, AMARTYA. "Isolation, Assurance, and the Social Rate of Discount," *Quart. J. Econ.*, Feb. 1967, 81(1), pp. 112-24.
- _____. *Collective choice and social welfare*. San Francisco: Holden Day, 1970a.
- _____. "The Impossibility of a Paretian Liberal," *J. Polit. Econ.*, Jan./Feb. 1970b, 78(1), pp. 152-57; reprinted in SEN 1982a, pp. 285-90.
- _____. "Choice Functions and Revealed Prefer-

- ence," *Rev. Econ. Stud.*, July 1971, 38(115), pp. 307-17.
- _____. "Behaviour and the Concept of Preference," *Economica N.S.*, Aug. 1973a, 40(159), pp. 241-59.
- _____. *On economic inequality*. Oxford: Clarendon Press, 1973b.
- _____. "Liberty, Unanimity and Rights," *Economica*, Aug. 1976a, 43(171), pp. 217-45.
- _____. "Real National Income," *Rev. Econ. Stud.*, Feb. 1976b, 43(1), pp. 19-39; printed in SEN 1982a, pp. 388-415.
- _____. "Rational Fools," *Philosophy and Public Affairs*, Summer 1977a, 6(4), pp. 317-44; reprinted in SEN 1982a, pp. 84-106.
- _____. "On Weights and Measures: Informational Constraints in Social Welfare Analysis," *Econometrica*, Oct. 1977b, 45(7), pp. 1539-72; reprinted in SEN 1982a, pp. 226-63.
- _____. "Social Choice Theory: A Re-examination," *Econometrica*, Jan. 1977c, 45(1), pp. 53-89; reprinted in SEN 1982a, pp. 158-200.
- _____. "Utilitarianism and Welfarism," *J. Philosophy*, Sept. 1979a, 76(9), pp. 463-89.
- _____. "The Welfare Basis of Real Income Comparisons," *J. Econ. Lit.*, Mar. 1979b, 17(1), pp. 1-45.
- _____. "Interpersonal Comparisons of Welfare," in *Economics and human welfare: Essays in honor of Tibor Scitovsky*. Ed.: MICHAEL BOSKIN. NY: Academic Press, 1979c, pp. 183-201; reprinted in SEN 1982a, pp. 264-81.
- _____. *Choice, welfare and measurement*. Cambridge, MA: MIT Press, 1982a.
- _____. "Rights and Agency," *Philosophy and Public Affairs*, Winter 1982b, 11(1), pp. 3-39.
- _____. "Evaluator Relativity and Consequential Evaluation," *Philosophy and Public Affairs*, Spring 1983a, 12(2), pp. 113-32.
- _____. "Liberty and Social Choice," *J. Philosophy*, Jan. 1983b, 80(1), pp. 5-28.
- _____. *Commodities and capabilities*. Vol. 7. Professor Dr. P. Hennipman lectures in economics, theory, institutions, policy. Amsterdam: North-Holland, 1985a.
- _____. "Rationality and Uncertainty," *Theory and Decision*, Mar. 1985b, 18(2), pp. 109-28.
- _____. "Well-being, Agency, and Freedom: The Dewey Lectures 1984," *J. Philosophy*, Apr. 1985c, 82(4), pp. 169-203.
- _____. "Foundations of Social Choice Theory: An Epilogue," in *Foundations of social choice theory*. Eds.: JON ELSTER AND AANUND HYLLAND. Cambridge: Cambridge U. Press, 1986a, pp. 213-48.
- _____. "Information and Invariance in Normative Choice," in *Social choice and public decision making: Essays in Honor of K. J. Arrow*. Vol 1. Eds.: WALTER HELLER, ROSS M. STARR, AND DAVID STARRETT. Cambridge: Cambridge U. Press, 1986b, pp. 29-55.
- _____. *On ethics and economics*. Oxford: Blackwell, 1987a.
- _____. "The Standard of Living: Lecture I, Concepts and Critiques," in SEN ET AL. 1987(b), pp. 1-19.
- _____. "The Standard of Living: Lecture II, Lives and Capabilities," in SEN ET AL. 1987(c), pp. 30-38.
- _____. "Gender and Cooperative Conflicts," in *Persistent inequalities*. Ed.: IRENE TINKER. London: Oxford U. Press, 1990, pp. 123-49.
- _____. *Inequality reexamined*. Cambridge, MA: Harvard U. Press, 1992.
- _____. "Minimal Liberty," *Economica*, forthcoming.
- SEN, AMARTYA ET AL. *The standard of living*. Ed.: GEOFFREY HAWTHORN. Cambridge: Cambridge U. Press, 1987.
- SEN, AMARTYA AND WILLIAMS, BERNARD, eds. *Utilitarianism and beyond*. Cambridge: Cambridge U. Press, 1982.
- SIDGWICK, HENRY. *The methods of ethics*. 6th ed. London: Macmillan, 1901.
- SINGER, PETER. "Altruism and commerce: A Defense of Titmuss against Arrow," *Philosophy and Public Affairs*, Spring 1973, 2(3), pp. 312-20.
- _____. "Freedoms and Utilities in the Distribution of Health Care," in DWORKIN, BERMANT, AND BROWN, eds. 1977, pp. 149-73.
- SMITH, ADAM. *The theory of moral sentiments*. Eds.: D. D. RAPHAEL AND A. L. MACFIE. Oxford: Clarendon Press, 1976.
- _____. *An inquiry into the nature and causes of the wealth of nations*. Eds.: R. H. CAMPBELL AND A. S. SKINNER [1776]; Oxford: Clarendon Press, 1976.
- SOLOW, ROBERT. "Wage Bargaining and Unemployment," *Amer. Econ. Rev.*, Dec. 1981, 71(5), pp. 898-908.
- _____. *The labor market as a social institution*. Oxford: Basil Blackwell, 1990.
- STEWART, HAMISH. "Taking Goals Seriously: A Reconsideration of Rationality in Economics." Ph.D. dissertation, Harvard U., 1989.
- STIGLER, GEORGE. "Wealth and Possibly Liberty," *J. Legal Stud.*, June 1978.
- STRASNICK, STEVEN. "Social Choice and the Derivation of Rawls's Difference Principle," *J. Philosophy*, Feb. 1976, 73(4), pp. 85-99.
- _____. "Neo-Utilitarian Ethics and the Ordinal Representation Assumption," in *Philosophy in economics: Papers deriving from and related to a workshop of testability and explanation in economics*. Ed.: JOSEPH C. PITT. Boston: D. Reidel, 1981, pp. 63-92.
- STREETEN, PAUL. "Appendix: Recent Controversies," in *The political element in the development of economic theory*. By GUNNAR MYRDAL, translated by PAUL STREETEN. London: Routledge & Kegan Paul, 1953, pp. 208-17.
- STROTZ, ROBERT. "Myopia and Inconsistency in Dynamic Utility Maximization," *Rev. Econ. Stud.*, 1956, 23(3), pp. 165-80.
- SUGDEN, ROBERT. "Liberty, Preference, and Choice," *Econ. Philos.*, Oct. 1985, pp. 213-29.
- _____. *The economics of rights, co-operation, and welfare*. NY: Blackwell, 1986.
- _____. "Spontaneous Order," *J. Econ. Perspectives*, Fall 1989, 3(4), pp. 85-97.

- _____. "Contractarianism and Norms," *Ethics*, July 1990, 100(4), pp. 768–86.
- SUZUMURA, KOTARO. *Rational choice, collective decisions, and social welfare*. NY: Cambridge U. Press, 1983.
- TAYLOR, CHARLES. *Hegel and modern society*. Cambridge U. Press, 1979.
- _____. *Sources of the self: The making of the modern identity*. Cambridge, MA: Harvard U. Press, 1989.
- TAYLOR, MICHAEL C. *The possibility of cooperation*. NY: Cambridge U. Press, 1987.
- THOMSON, GARRETT. *Needs*. London: Routledge, 1987.
- THOMSON, WILLIAM AND LENSBERG, TERJE. *Axiomatic theory of bargaining with a variable number of agents*. NY: Cambridge U. Press, 1989.
- THUROW, LESTER C. "Government Expenditures: Cash or In-Kind Aid," in GERALD DWORKIN, BERMAN, BROWN, eds. 1977, pp. 85–106.
- TITMUSS, RICHARD. *The gift relationship: From human blood to social policy*. NY: Random House, 1971.
- TVERSKY, AMOS AND KAHNEMAN, DANIEL. "Prospect Theory: An Analysis of Decision under Risk," *Econometrica*, Mar. 1979, 47(2), pp. 263–91.
- VANBERG, VIKTOR. "Spontaneous Market Order and Social Rules: A Critical Examination of F. A. Hayek's Theory of Cultural Evolution," *Econ. Philos.*, Apr. 1986, 2(1), pp. 75–100.
- VARIAN, HAL. "Equity, Envy, and Efficiency," *J. Econ. Theory*, Sept. 1974, 9(1), pp. 63–91.
- _____. "Distributive Justice, Welfare, Economics, and the Theory of Fairness," *Philosophy and Public Affairs*, Spring 1975, 4(3), pp. 223–47.
- _____. "Dworkin on Equality of Resources," *Econ. Philos.*, Apr. 1985, 1(1), pp. 110–25.
- VICKREY, WILLIAM. "Measuring Marginal Utility by Reactions to Risk," *Econometrica*, Oct. 1945, 13(4), pp. 319–33.
- _____. "Utility, Strategy, and Social Decision Rules," *Quart. J. Econ.*, Nov. 1960, 74(4), pp. 507–35.
- WALDRON, JEREMY, ed. *Theories of rights*. Oxford: Oxford U. Press, 1984.
- WALZER, MICHAEL. *Spheres of justice*. NY: Basic Books, 1983.
- WARR, PETER G. AND WRIGHT, BRIAN D. "The Isolation Paradox and the Discount Rate for Benefit-Cost Analysis," *Quart. J. Econ.*, Feb. 1981, 96(1), pp. 129–45.
- WEIBULL, JÖRGEN. "Persistent Unemployment as Subgame Perfect Equilibrium." Seminar Paper No. 381 of the Institute for International Economic Studies, Stockholm, May 1987.
- WESTON, SAMUEL. "Toward a Better Understanding of the Postitive/Normative Distinction in Economics," *Econ. Philos.*, forthcoming.
- WEYMARK, JOHN. "A Reconsideration of the Harsanyi-Sen Debate on Utilitarianism," in ELSTER AND ROEMER, eds. 1991, pp. 255–320.
- WILLIAMS, BERNARD. "A Critique of Utilitarianism," in *Utilitarianism: For and against*. Eds.: J. J. C. SMART AND BERNARD WILLIAMS. Cambridge: Cambridge U. Press, 1973, pp. 77–150.
- _____. "The Standard of Living: Interests and Capabilities," in SEN ET AL. 1987, pp. 94–102.
- WILLIAMSON, OLIVER. "The Organization of Work: A Comparative Institutional Assessment," *J. Econ. Behav. Organ.*, 1980, 1(1), pp. 5–38.
- YAARI, MENACHEM AND BAR-HILLEL, MAJA. "On Dividing Justly," *Soc. Choice Welfare*, 1984, 1(1), pp. 1–24.