

COPYRIGHT NOTICE:

Russell Hardin: Indeterminacy and Society

is published by Princeton University Press and copyrighted, © 2003, by Princeton University Press. All rights reserved. No part of this book may be reproduced in any form by any electronic or mechanical means (including photocopying, recording, or information storage and retrieval) without permission in writing from the publisher, except for reading and browsing via the World Wide Web. Users are not permitted to mount this file on any network servers.

For COURSE PACK and other PERMISSIONS, refer to entry on previous page. For more information, send e-mail to permissions@pupress.princeton.edu

Chapter One

Indeterminacy

INDETERMINACY in contexts of strategic interaction—which is to say, in virtually all social contexts—is an issue that is constantly swept under the rug because it is often disruptive to pristine social theory. But the theory is fake: the indeterminacy is real. I wish here to address such indeterminacy, its implications for collective choice, and the ways in which it has been hidden from view or ignored in manifold theories, and some ways in which it has been well handled and even made central to theory. The effort to pretend indeterminacy away or to hide it from view pervades social theory of virtually every kind, from the most technical game-theoretic accounts of extremely fine points to moral theory from its beginnings. The issue is that the basic rationality that makes sense of or fits individual choice in the simplest contexts of choosing against nature does not readily generalize to contexts in which individuals are interacting with other individuals. The basic rationality that says more resources are preferable to less is indeterminate for the more complicated context, which is almost the whole of our lives outside the casino and the lottery.

Typically, the central task in strategic interactions is obtaining the best possible outcome for oneself. Unfortunately, in many social contexts I cannot simply act in a way that determines my own outcome. I can choose only a strategy, not an outcome. All that I determine with my strategy choice is some constraints on the possible array of outcomes I might get. To narrow this array to a single outcome requires action from you and perhaps many others. I commonly cannot know what is the best strategy choice for me to make unless I know what strategy choices others will make. But if all of us can know what all others are going to do, then it is not coherent to say that thereafter we can alter our choices in the light of that knowledge. This is the form of indeterminacy at issue in this book: *indeterminacy that results from strategic interaction*. Interactive choice as represented descriptively in game theory is often indeterminate for each individual chooser. For an individual chooser in a moment of choice, this indeterminacy is not a failure of reason by the chooser; the indeterminacy is in the world because it follows from the mismatch of the preferences of all those in the interaction of the moment.

Problems of a partially related kind were characterized by C. H. Waddington ([1960] 1967, 17), in a sadly neglected book, as *stochastic*. This word, which

means, roughly, “probabilistic,” comes from a Greek root that means “proceeding by guesswork” or, more literally, “skillful in aiming.” Stochastic problems are those for which, in a sense, nature might outsmart our choice of strategy so that we get an outcome very different from what we would have wanted, at least in some cases. As a remarkably clear and varied example of the problems at issue, I will often discuss problems of vaccination against a major disease. That range of problems has the attractive feature that it, in its simpler forms, should not be very controversial, either pragmatically, causally, or morally. Especially because it is often not morally controversial, it will be very useful in exemplifying the nature of stochastic policy problems, including many, such as nuclear deterrence (in which accidents could have happened), that are also troubled with strategic interaction. For stochastic problems of individual choice, we can readily reduce our choices to their expected values; and then we can select more rather than less. Stochastic collective choices or policies commonly entail losses for some and gains for others, so that we can choose unproblematically from expected value only if we do not know in advance who will be the losers and gainers.

To see the peculiarly stochastic nature of many collective choice problems that we face, consider the program of polio vaccination before it was eradicated as a disease in the wild in North America (that is, outside certain laboratory stores of the virus). The facts are roughly these. We vaccinate millions, including almost the entire population of children. Many of these would die and many would be permanently, even hideously, crippled if not vaccinated. Among those vaccinated, a very small number suffer serious cases of paralytic polio. There is no question that fewer are harmed by vaccinating than by not vaccinating the population. Our strategic action is to protect people, but among the outcomes that could follow from our action, we also harm some people, some of whom might never have got polio if not vaccinated or even if no one had been vaccinated.¹

When we choose an action or a policy, this is often the structure of it. We have some chance of doing harm and some chance of doing good in the unavoidable sense that in order to do something good we must risk doing something bad. Sometimes this is for reasons of the nature of the world, as in the case of vaccination. But at other times it is for reasons of the nature of strategic interaction. I choose, in a sense, a strategy, not an outcome. Then I get an outcome that is the result of the strategy choices of others in interaction with my choice.

These two classes of problems, strategic interaction and stochastic patterns of outcomes, have a common feature, which is the central issue of this book. They make indeterminate what an action or a policy is. In philosophical action theory, the actions are simple ones such as flipping a switch to turn on a light. In real life, our most important actions are not so simple. They are inherently interactions. We have reasons for taking our actions, but our reasons may not

finally be reflected in the results of our actions even if hope for specific results is our reason for our choice of actions.

In three contexts I argue that taking indeterminacy into account up front by making it an assumption helps us to analyze certain problems correctly or to resolve them successfully. In these cases, using theories that ignore the indeterminacy at issue can lead to clouded understandings, even wrong understandings of the relevant issues. One of these contexts is the hoary problem of the iterated prisoner's dilemma and what strategy is rational when playing it. The second is the real world prisoner's dilemma of nuclear deterrence policy that, one hopes, is now past. The third is the great classical problem of how we can justify institutional actions that violate honored principles. For example, public policy is often determined by a cost-benefit analysis, which entails interpersonal comparisons of utility. The people who do these policy analyses are commonly economists who eschew interpersonal comparisons as metaphysically meaningless. Such comparisons are one theoretical device for avoiding indeterminacy. Although they have been intellectually rejected on theoretical grounds, and seemingly rightly so, still they make eminently good sense on an account of their authorization that is grounded in indeterminacy. In all three of these contexts, by starting with the—correct—presumption of indeterminacy, we get to a better outcome than if we insist on imposing determinacy.

Note that the vaccination case, in which some are harmed while others are benefited, is only a partial or reduced analogue of the social interaction case in which depending on what you do, I may do very well or very badly. The problem is simplified in that one of the actors is nature, rather than a strategically manipulative agent with its own interests possibly in conflict with ours. Keeping this simplified problem in mind helps to make the more complex issues of strategic interaction between two or more manipulative, self-interested agents relatively clear. The causal analysis of stochastic problems such as vaccination policy is unlikely to be controversial, whereas any account of a strategic interaction that stipulates what each of the parties will do or ought to do is likely to be controversial. Indeed, that is the central fact that provokes this work.

STRATEGIC INTERACTION

It is a correct assessment of rationality in social contexts that it is ill defined and often indeterminate. If this is true, then any instruction on what it is rational to do should not be based on the (wrong) assumption of determinacy. Assuming that the world of social choice is indeterminate rather than determinate would lead one to make different decisions in many contexts. Determinacy can be both disabling and enabling, depending on the nature of the decisions at stake. The indeterminacy of a principle of rational choice sometimes plays

into the indeterminacy of knowledge, but I wish to address problems of knowledge primarily as they affect rational choice through strategic or rational indeterminacy, not as they affect ordinary decision making through epistemological indeterminacy.

Epistemological indeterminacy from causal ignorance—for example, from inadequate theory or inadequate knowledge—is a major problem in its own right, but it is not the focus here. Hence, I am not concerned with the disruptive possibilities of unintended consequences (which are a major problem in innovation and in public policy), even insofar as they result from complex interactions.

Strategic or rational indeterminacy, as in current theory, is partly the product of the ordinal revolution in economic and choice theory. That revolution has swept up economics and utilitarianism and has helped spawn rational choice theory through the ordinal theories of Kenneth Arrow ([1951] 1963) and Joseph Schumpeter ([1942] 1950). The problem of such indeterminacy arises from simple aggregation of interests, and therefore it is pervasive in neoclassical economics as well as in ordinal utilitarianism or welfarism. It is pervasive *because our choices have social (or interactive) contexts*. Arrow demonstrated this indeterminacy in principle already in one of the founding works of social choice. It is instructive that he discovered the indeterminacy while trying to find a determinate solution to collective aggregation of ordinal preferences. Unlike most theorists, however, he did not quail from the discovery but made it the center piece of his Impossibility Theorem (Arrow 1983, 1–4).

Because there is collective indeterminacy, there is indeterminacy in individual choice in contexts of strategic interaction. These are, in a sense, contexts of aggregation of interests, even though there may be substantial conflict over how to aggregate and those in interaction need not be concerned at all with the aggregate but only with personal outcomes. In such interactions, we may treat each other as merely part of the furniture of the universe with which we have to deal, so that we have no direct concern with the aggregate outcome, only with our own. As John Rawls (1999, 112 [1971, 128]) supposes, we are mutually disinterested. We should finally see such indeterminacy not as an anomaly but as the normal state of affairs, on which theory should build. Theory that runs afoul of such indeterminacy is often foul theory.

A quick survey of contexts of strategic interaction in which indeterminacy has played an important role and in which theorists have attempted to get around it or to deal with it would include at least the following seven (each of these will be discussed more fully in later chapters).

In game theory, John Harsanyi simply stipulates that a solution theory must be determinate despite the fact that adopting his determinacy principle makes no sense as an optimizing move. His move comes from nowhere, as though somehow it is irrational to live with indeterminacy (in which case it is irrational to live). It appears to be a response to the oddity of the prisoner's dilemma game when this game is iterated for a fixed number of plays (see chapter 2).

That game is a pervasive part of life because it is essentially the structure of exchange. Any economist's theory of rationality must be able to handle that game. One might even say that the analysis of that game should come before almost anything else in the economist's world.

Equilibrium theory in economics is fundamentally indeterminate if there is a coordination problem. There is a coordination problem whenever there is more than one coordination equilibrium. In any whole economy, to which general equilibrium is meant to apply, there are apt to be numerous coordination problems in the abstract (see chapter 2).

Thomas Hobbes attempted to trick up determinacy in his theory of the creation of a sovereign, although he needed no trickery in his selection of any extant sovereign as determinately preferable to putting any alternative in place by rebellion (see chapter 3).

Jeremy Bentham imposed determinacy in his version of utilitarianism by supposing that utilities are comparable and additive across persons, so that in a comparison of various states of the universe, we could supposedly add up the utilities and immediately discover which state has the highest utility (see chapter 4).

Ronald Coase, with his Coase Theorem, may have made the cleverest move to overcome the problem of indeterminacy in an ordinal world by using cardinal prices to resolve the choice of what to produce (see chapter 5), although his resolution of this problem still leaves open the question of how to share the gains from production among the owners of the relevant productive assets.

A standard move in much of moral theory is to reduce the inordinate multiplicity of possible problems of individual choice by introducing a set of rules that grossly simplifies the choices we must make (see chapter 6). Such moral theory is now called deontology.

In his theory of justice, John Rawls achieves the appearance of determinacy in the abstract with his difference principle, but under that appearance there is a morass of indeterminacy in his category of primary goods that, if taken seriously, negates much of the seeming simplicity and appeal of his theory (see chapter 7). Nevertheless, far more clearly than most of the theorists considered here, he recognizes the centrality of the problem of indeterminacy in social choice and very cleverly attempts to overcome it.

We may ex ante create institutions that make decisions on principles that we would not have been able to use directly. For example, it may be mutually advantageous ex ante for us to have an institution use cost-benefit analysis, with its attendant interpersonal comparisons, even though we might not be able to give a mutual-advantage defense of such an analysis in any particular instance of its application (see chapter 8).

The responses to these contexts include three failures of theory: to ignore the problem and suppose that choice theory is determinate (as in Harsanyi's game theory and in equilibrium theory); to cardinalize the values at issue so

that they can then be added up in various states and the highest value can be selected (as in Bentham's utilitarianism); and to adopt very limited but relatively precise rules or principles for behavior that cover some limited range of things and to exempt other things from coverage (as in deontological ethics).

There are three pragmatic responses that are variously effective: to simplify the problem so that ordinal resolution is relatively determinate (as in the moves of Hobbes and Rawls, as discussed in chapters 3 and 7); to keep everything ordinal and noncomparable up to the point of comparing the cardinal market values of what is produced (as in Coase's theorem in chapter 5); and to shift the burden of choice to an institution to achieve mechanical determinacy (as discussed in chapter 8).

Finally, there is also the possibility of accepting indeterminacy and resolving issues by making indeterminacy an assumption or conclusion of the analysis, as in Arrow's theorem, rather than a problem to be ignored or attended to later.

Each of these responses yields determinate solutions to problems up to the limit of the device. The first three devices, however, block out of view the fundamental indeterminacy of the choice or assessment problem. And the three pragmatic devices may obscure the underlying indeterminacy. I will discuss these six devices in subsequent chapters. Shifting the burden of choice *ex ante* to an institution is in part the resolution of Hobbes and Rawls, but it is more substantially the way we handle public policies in the face of the other, generally less workable resolutions in the list. The last way of dealing with indeterminacy—simply to face it and incorporate it into analysis—is, so far, not very common. I think it yields a correct analysis of social choice in an ordinal world in Arrow's Impossibility Theorem, a finally correct analysis of how to play in iterated prisoner's dilemma (which is a good model of much of the life of exchange and cooperation), a credible and effective account of how to handle such issues as nuclear arms control, and the richest and most compelling vision of social order that we know. It also leads Rawls to his difference principle for handling plural values. It additionally fits many other problems and theories that will not be addressed here, including general solution theories for games,² the size or minimum-winning coalition theory of William Riker (1962; Hardin 1976), chaos theory in electoral choice (Mueller 1989), and many others.

The varied ways of dealing with indeterminacy have distinctively different realms of application, as briefly noted above and as I will discuss further in later chapters. For example, Hobbes's grand simplification of the problem of social order allowed him to conclude that any stable government is better than none and that loyalty to any extant government is better than any attempt to put a better one in its place. But it also works in some cases in which simplification of the interests at stake is not even necessary. Coase's resolution is of marginal problems of allocating resources for production. It arises as a problem only after Hobbes's or some other resolution of the general problem of social

order has been achieved. Cardinalization with interpersonal comparisons of welfare would work at any level, from foundational to marginal, if it could be made intelligible, as sometimes perhaps it can be. At the very least, we frequently act as though it makes sense, often in fundamentally important contexts, such as in many public policy decisions.

The first three of these devices are efforts to trick out as much as possible from what indeterminacy there may be. When any of them works, it is fine, although many people, especially including economists, reject any appeal to interpersonal comparison of welfare as meaningless. But the devices do not always work very well, and then we are left with trying to deal with indeterminacy or trying to trick ourselves, rather than the world, into believing, for example, that a limited set of moral rules can be adequate for at least morality. The responses are of quite different kinds, and there have probably been other devices of significance. Their sophistication and variety suggest how pervasive and varied the problem of indeterminacy is.

By now, one might suppose we would have recognized that indeterminacy is part of the nature of our problems in various contexts. Tricking or fencing it out of the picture is commonly misguided and will not help us resolve or even understand many of our individual and social choice problems, as Arrow clearly understood. Instead of imposing determinate principles on an indeterminate world in some contexts, we should often seek principles that build on the indeterminacy we must master if we are to do well. If we do this, we will most likely find that the principles we apply will be incomplete. They will apply to choices over some ranges of possibilities and not over others.

For many choice contexts, the principle we might adopt is the pragmatic principle of melioration rather than maximization, which is inherently undefinable for many contexts. I will take melioration to be modal advantage, which is expected advantage to all *ex ante*, although it may commonly happen that not all gain *ex post*. For social choice and for moral judgment of aggregate outcomes, the principle for which I will argue is, *when it is not indeterminate*, mutual advantage. This principle will often make sense *ex ante* although not *ex post*, as will be exemplified in a discussion of vaccination (chapter 3). In actual life, especially in politics, we will more likely see a limited version of it that we might call sociological mutual advantage. Our political arrangements will serve the mutual advantage of groups that have the power to block alternatives and may substantially ignore groups without such power (Hardin 1999d, chap. 1 and *passim*). Sociological mutual advantage, however, lacks the normative weight of fully inclusive mutual advantage.

When mutual advantage is not decisive and we nevertheless have to decide what to do, we may have recourse to an aggregate variant of melioration in which some interpersonal comparisons might be made. *This is itself a mutual advantage move ex ante*. That is to say, we know in advance that mutual advan-

tage in case-by-case collective decisions will not work. We therefore need a principle for handling those cases. Adopting a relatively loose principle of melioration for handling such cases when they arise can be mutually advantageous even though its specific applications will not be. Even more commonly, we create an institution that will then apply such a meliorative principle, so that we get mechanical determinacy.

The claim for indeterminacy here is not merely another silly metaphor on quantum mechanics. Interactive choice and aggregate valuation just are sometimes indeterminate. This is not an analog or metaphorical extension of indeterminacy in physics. Social indeterminacy is a problem of set-theoretic choice theory rather than of physical possibilities. *The central problem is indeterminacy of reason in the face of strategic interaction.* Terms in the family *rational* often seem to lose their clarity and supposed definitiveness in such contexts. In fact, they are not univocally definitive. Beyond contexts in which basic rationality—to prefer more to less value when nothing else is at stake in the choice—suffices for choice, if we wish to have determinate theory, we must bring in many additional principles that are often ad hoc and that are not as compelling as basic rationality. Stochastic problems are simpler than those of complex strategic interaction. In them, I or we commonly get one from a probabilistic array of possibilities and we cannot narrow the choice to the best possibility in the array, although it might turn out to be the outcome we get. In such choices, we are, in a sense, choosing against nature, who plays her own strategy.

Indeterminacy in strategic interaction has been clearly recognized in certain contexts for more than two centuries. In the era of the French Revolution, the marquis de Condorcet recognized that in the context of majority voting, there can be cyclic majorities in which, say, candidate *A* defeats candidate *B* by a majority, *B* defeats *C* by a different majority, and *C* defeats *A* by yet another majority. Hence, the principle of majority choice is indeterminate. Moreover, at least since Hobbes the problem of indeterminacy has troubled many social theorists, most of whom have attempted to dodge it. Often, the greater part of wisdom is finally to recognize its pervasiveness and to deal with it by grounding our theories in it, not by building pristine, determinate theories and then trying to force social reality to fit these, or even criticizing it for failing to fit.

We often also face causal indeterminacy, which will come into discussion here insofar as it plays a strong role in defining the possible ways of dealing with indeterminacy in strategic or rational choice. Causal indeterminacy might often be merely a lack of adequate understanding of causal relations, so that it might be corrected. Rational or social indeterminacy is not merely a lack of understanding, and it cannot be corrected, although we might find pragmatic dodges or tricks to help us master the problems it creates.

ORDINALISM

The principal reason for social indeterminacy is the pluralism of values and evaluators typically involved in social choice. A brief account of ordinalism and its intellectual origins is therefore useful. The clear and explicit understanding of social choice as an ordinal problem began with Vilfredo Pareto's insights that lie behind what are now called the Pareto criteria: Pareto efficiency or optimality and Pareto superiority. A state of affairs Q is Pareto superior to another state P if at least one person is better off in Q than in P and no one is worse off in Q than in P . And a state of affairs Q is Pareto efficient or optimal if there is no other state that is Pareto superior to it. The latter condition implies that any move would make at least one person worse off or would make no one better off.

It seems plausible that, implicitly, ordinal valuations were commonly assumed in many writings long before Pareto. For example, Thomas Hobbes, David Hume, and Adam Smith were generally ordinalists. But the clear focus on ordinalism and the attempt to analyze its implications came as a response to the suppositions, overt in Bentham and often seemingly tacit before Bentham, that utility is cardinal and that utility across persons can be added. That view dominated much of nineteenth-century thought. But cardinalism was a mistaken invention of high theory. The task that Pareto faced was to make systematic sense of ordinalism. That task was finally carried to fruition in the so-called ordinal revolution in economics in the 1930s (see Samuelson 1974). The implausibility of cardinal, additive utility was already problematic for understanding how prices could vary. The marginal revolution of the nineteenth century was half of the mastery of that problem; the mastery was completed in the ordinal revolution.

The ordinal Pareto criteria are commonly treated as though they were rationally not problematic. But they are subject to strategic problems in practice if they are taken to imply recommendations for action. Unfortunately, the Pareto criteria are about valuations of the states of affairs, and the valuations do not guarantee the actions—exchanges—that would bring us to those states. Moves to Pareto superior states are often thought to be rational because everyone will consent to them because there will be no losers, only gainers, from such moves. But some might not consent. Why? Because any move will not merely determine a present improvement but also will determine what states will be reachable from the newly reached state. Doing especially well—compared to others—on the first move raises the floor from which one maneuvers for the next and subsequent moves. Doing badly on the first move, while nevertheless benefiting at least somewhat in comparison to one's status quo point, lowers the ceiling to which one can aspire on later moves. Figure 1.1 shows this point more clearly.

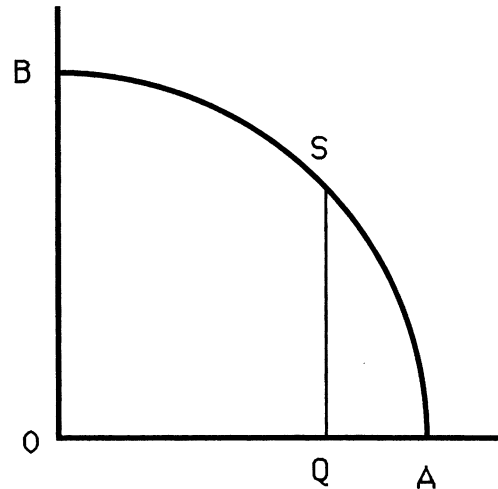


Fig 1.1. Pareto frontier

Figure 1.1 represents a distribution of resources between A and B beginning from the initial or status quo distribution at the origin, O . The Pareto frontier represents the set of distributions that are Pareto optimal and also Pareto superior to the status quo. We can imagine that A and B have holdings of various goods at the status quo distribution and that, through exchange, they can both be made better off. When they reach a stage at which no further voluntary exchanges can be made, they will be at some point on the frontier, represented by the arc from A to B . On the frontier, any move or exchange that would make A better off would make B worse off. Any point between the origin and the frontier is accessible from the origin with a Pareto-superior move. For example, the move from the origin to the point Q (inside the frontier) makes A better off without affecting B 's welfare. A move from Q back to O , however, is not a Pareto-superior move because it reduces the welfare of A . But note what a move from O to Q does. It effectively defines a new Pareto frontier that includes all points of the original frontier (from A to S) that are not to the left of Q . All of the points that have been eliminated (all those from B to S) would be preferred by B to any of the points that remain, while all of those that remain would be preferred by A to any that have been eliminated. In a meaningful sense, A has gained while B has lost in the move from O to Q . B 's loss could be called an opportunity cost.

Why then should we say it is rational for B to consent to the Pareto-superior move from O to Q ? It is self-evidently rational for B to consent to the move *only if* it excludes no points that B would prefer to the remaining Pareto-superior points. This will be true if the move from O to Q is in the only direction

in which it is possible to move. But this just means that the Pareto frontier must not be a curve as drawn in figure 1, but must be a point directly to the right of the origin. If the Pareto frontier is as drawn in figure 1, *B* cannot rationally consent to any move without first considering its effect on an eventual Pareto-optimal distribution. Hence, Pareto superiority is not a criterion for rational choice.

If I am narrowly rational, I can be indifferent to what you gain from a trade from me only if it does not potentially restrict what I can gain from further trades with you. For *A* and *B* in the situation of figure 1, however, every Pareto-superior move involves opportunity costs for at least one of the two and often for both of them. That is to say, the move eliminates some attractive opportunities from the field of subsequent choice.³ Hence, one cannot view the criterion of Pareto superiority as a principle of rational choice in general. The criterion is rationally objectionable because it presumes consent in situations in which conflict can arise over potential allocations. Because my future opportunities may depend on how we allocate things now, my interests now depend on my future prospects.

The Pareto criteria are commonly indeterminate in judging pairs of outcomes even if these problems do not afflict them. For example, any two points on the frontier are Pareto noncomparable. We cannot say that either of these states of affairs is Pareto superior to the other. If opportunity costs are taken into account, we may further not be able to say that any pair of states of affairs is Pareto comparable, with one state superior to the other. Alas, therefore, the Pareto principles reintroduce indeterminacy if they are taken to recommend action.

Pareto combined Hobbes's normative principle of mutual advantage with marginalist concern. Indeed, Pareto ([1927] 1971, 47–51) formulated his principles to avoid interpersonal comparisons and attendant moral judgments. Hence, Pareto was Hobbesian in his motivations, at least in part. The criteria were introduced by Pareto not for recommending individual action but for making ordinal value judgments about states of affairs. They might therefore be used by a policy maker.

The Pareto principles are static principles about the distribution of those goods we already have. They are not dynamic or production-oriented. In this respect, they might seem to be an aberration, as is Benthamite additive utility if it is to be done with precision. But the Paretian analysis of static efficiency accounts for a real issue and is therefore not merely an aberration. It may nevertheless often be largely beside the point for our purposes, although it comes back in through Coasean efficiency in law and economics. The indeterminacy of the notion of Pareto improvement is in the allocation of some surplus available to us beyond the status quo. Most allocations of the surplus might make every one of us better off and would, therefore, be mutually advantageous.⁴

MUTUAL ADVANTAGE: THE COLLECTIVE IMPLICATION
OF SELF-INTEREST

During the century or two ending about 1900, microeconomics and utilitarianism developed together. At the beginning of the twentieth century G. E. Moore was the first major utilitarian philosopher who did not also write on economics; indeed, economists might not be surprised to learn that he mangled value theory and therewith utilitarianism. His value theory reverted to the crude notion that some value inheres in objects independently of anyone's use of or pleasure in those objects (Moore 1903, 84). A variant of this notion lies behind the labor theory of value, according to which the value of the object is the quantity of total labor time for producing it—the Salieri theory of value.⁵ John Stuart Mill and Henry Sidgwick, the great nineteenth-century utilitarians, both wrote treatises on economics, and much of what they say about value theory is grounded in their economic understandings. Moore returned speculation on value to the Platonic mode of pure reason, so called.

It is ironic that many of the outstanding problems in economic value theory were resolved in Moore's time by Pareto, as will be discussed in chapter 3, and the exponents of the ordinal revolution who followed his lead in the 1930s. Moore was evidently oblivious of these developments. Major economists of Moore's time were avowedly utilitarian (Edgeworth 1881; Pigou [1920] 1932), yet intellectually they had little in common with Moore. Moore's one great insight into value theory—that the value of a whole is not a simple addition of the values of its parts—is one of the problems that stimulated the ordinal revolution and that was resolved by it (see Hardin 1988, 7; Moore 1906, 28). That insight was not entirely original with Moore. Mill had earlier referred to the analogous issue in objective causal contexts as a matter of "chemical" combinations, because when two chemicals are brought together, the result is commonly a chemical reaction that produces a compound with none of the distinctive properties of the original chemicals—as combining the two gases hydrogen and oxygen produces water. So too, in value theory, bringing two things together may produce a combined value that is unrelated to the sum of the values of the two things taken separately.

Many (perhaps most) economists in the Anglo-Saxon tradition have continued to be utilitarian insofar as they have a concern with moral theory or normative evaluation. Unfortunately, at the time of Moore, utilitarianism in philosophy separated from economics and therefore from developments in economic value theory—just when these could have helped to reconstruct or continue the intellectual development of utilitarianism. The two traditions have been brought back together most extensively since then in the contemporary movement of law and economics. This rejoining recalls the early origins of the general utilitarian justification of government in the theory of Hobbes. The

chief difference is that Hobbes's concern was foundational, whereas the concern of law and economics is marginalist. Hobbes was concerned with establishing social order at all; law and economics focuses on rules for legal allocations in an already well ordered society. In both law and economics and in Hobbes's theory of the sovereign, a principal focus is on normative justification. And in both, the basic principle of normative justification is self-interest somehow generalized to the collective level, as discussed further below.

The distinctive unity of the visions of those I will canvass in the development from Hobbes to Coase is that it is welfarist. But Hobbes, Pareto, and Coase are not part of the Benthamite classical utilitarian aberration in interpersonally additive welfarism. It is therefore perhaps less misleading here to speak of welfare rather than of utility, although the contemporary notion of utility is manifold in its range of meanings. The chief reason for speaking of welfare rather than utility is that the language of welfare is different from that of utility. Typically we do not want *more* welfare, although we often want greater welfare or a higher level of welfare. Welfare is not a cardinal value made up of smaller bits of welfare, as utility is sometimes assumed to be. The language of welfare is typically ordinalist. I will generally refer to utility only in discussing Bentham's views and will otherwise speak of welfare or well-being.

The normative foundations of the formulations of Hobbes, Vilfredo Pareto, and Ronald Coase are essentially the same. Hobbes emphasized that all we have are individual values (essentially the values of self-interest: survival and welfare) and that individuals can be motivated by resolutions of their collective problem only if these speak to their interests. Pareto, writing after a century of Benthamite additive utilitarianism, asserted that we do not know what aggregation of utility across individuals means, that it is a metaphysical notion. Both Hobbes and Pareto concluded that we can ground a motivational theory only in disaggregated individual values. The only collective value that can be directly imputed from these is mutual advantage, in which all gain (Hobbes's usual assumption) or at least none loses while at least one gains (Pareto's assumption).

Coase ([1960] 1988) is less concerned to state his value position than were Hobbes and Pareto, but his position also seems clearly to suppose that collective values are reducible to individual values or that the only values of concern in his accounts are individual values. An example of this move is in Coase's discussion of the negotiations between a rancher and a farmer on the optimizing of returns from their joint enterprises. The total net profits from the farmer's crops and the rancher's cattle are a function of their sales value less the costs of raising them. This total is essentially a cardinal value in, say, dollars. Suppose these profits could be increased by letting the cattle roam over part of the farmer's land, destroying some of her crops, but that the farmer has the right to fence her land against the cattle. The two then have an interest in striking a deal that allows the cattle to roam over part of the farmer's land (Coase [1960]

1988, 95, 99). In this deal, part of the extra profits the rancher makes from the farmer's not fencing his cattle off her land go to the farmer as compensation for lost profits from her damaged crops, and the two split the remaining extra profits. Hence, each can be made ordinally better off by making the deal. In rough terms, this example generalizes as Coase's theorem.

For ordinalists such as Hobbes, Pareto, and Coase, we may say that *mutual advantage is the collective implication of self-interest* because to say that an outcome of our choices is mutually advantageous is to say that it serves the interest of each and every one of us. One could say that, in this view, collective value is emergent; it is merely what individuals want. But one could also say that this is what the value is: to let individual values prevail. To speak of collective value in any other sense is to import some additional notion of value into the discussion beyond the interests of individuals. Hobbes may have been constitutionally oblivious of any such additional notions of value; Pareto evidently believed them perverse.

Incidentally, mutual advantage is the only plausible collective analog of self-interest. Consider a cardinally additive measure, for example. More utility to you can compensate for less utility to me in such an additive measure. Of course, it cannot be my interest for you to benefit at my expense. Hence, a cardinally additive measure is not a collective analog of individual-level self-interest.

CONCLUDING REMARKS

Almost all of the discussions here are welfarist. The indeterminacies have to do with interests and with their aggregation or balancing in some way. In chapter 2, the concern is with the definition of interests. In subsequent chapters, it is more generally with the aggregation or balancing of interests. Sometimes resources come into play, as in discussions of Posner's wealth maximization (chapter 4), Coase's theorem (chapter 5) and Rawls's theory of justice (chapter 7). Sometimes, resources have been invoked as a genuine alternative to welfare for various conceptual reasons, as in their use by Posner and Rawls and in Amartya Sen's (1985, 1999) arguments for capacities as the core concern of political theory. Sometimes they are merely a useful device for dealing with welfare, as in Coase's theorem. In the discussion of moral rules as a device to ignore indeterminacies (chapter 6), welfare and resources are not generally at issue.

Few technical problems in rational choice have been hashed out and fought over as intensely as the analysis of how to play in an iterated prisoner's dilemma (chapter 2). And no problem in all the millennial history of political philosophy has been more central and debated than how to justify government and its actions (chapters 3, 7, and 8). Both these problems are substantially clarified by recognizing that they are indeterminate in important ways, so that our theories for handling them must be grounded in indeterminacy.

The critical success of some theories historically has probably depended substantially on papering over the indeterminacy that undercuts them, as in the case of Locke's contractarianism, more recent variants of arguments from reasonable agreement and deliberative democracy, and Rawls's theory of distributive justice. Despite its flaccidity, Locke's contractarianism has commonly been taken to be a better account of the justification of government than has Hobbes's theory. This judgment is badly wrong. Hobbes achieves the astonishing success of founding government in an account from self-interest, or rather from its collective implication in mutual advantage. Locke's theory leaves us dependent on a normative commitment that is not credible.

There are many related arguments along the way. For example, if rational choice were determinate, then it would tend to produce equilibrium outcomes. But if rationality is indeterminate, equilibrium may also be indeterminate, as I argue it is, if properly conceived, in the iterated prisoner's dilemma (chapter 2). If we do reach equilibrium outcomes, we may generally expect to be stuck with them. But in many contexts there will in general be no reason to expect to reach equilibrium because there will be none. Daily life and political theory lack equilibriums, and we should probably be glad of that—at least if we are welfarist, as we must be to some extent no matter how much we might assert that welfarism is a wrong or bad moral theory. To paraphrase Keynes, at equilibrium, we are all dead—or at least we have stopped all production.

Although I will not argue this claim extensively here, if welfarism is a bad moral theory, then it is also a bad pragmatic theory. But to suppose it a bad pragmatic theory is so utterly implausible that one cannot sensibly hold that it is a bad moral theory. At the very least, it must be a large part of any good moral theory. This is most conspicuously true, perhaps, in the context of designing institutions for our social life.

My purpose in the discussions of the following chapters is not specifically to explicate the relevant theories of, for example, Hobbes, Kant, Bentham, Coase, Rawls, or Posner. I merely address the problems and successes their theories may have with indeterminacy, because it is the significance of indeterminacy in society and in strategic interactions that I wish to understand. The strengths of some of these theories are arguably more impressive than their weaknesses, but I have deliberately focused on one aspect of their weaknesses.