4 The normative core of rational choice theory

Russell Hardin*

Introduction

Although it is commonly called a positive theory – to imply that it is purely descriptive and without value assumptions – rational choice theory is typically grounded in a powerful and simple value theory, from which many of its most compelling results follow. In actual applications, the word “rational” in rational choice theory is typically not a merely formal term. It is also a substantive term that refers to particular values, essentially welfare values. It may be true, as some rational choice theorists insist, that the theory could be applied to actors whose values are other than welfare and, especially, other than their own welfare. But the remarkable success of rational choice explanations turns on the surprising power they have when they are grounded almost exclusively in the actors’ own welfares.

To this extent, rational choice theory is, practically, a two-fold theory that says what values govern individuals’ choices and what the result of those choices will be. Often, the result is collective rather than merely individual, although rational choice explanations are strictly at the individual level and can govern individual actions outside social interactions as well as within such interactions. In the discussion here, however, I will focus on rational choice explanations of social-political interactions, including large group interactions.

A common complaint against rational choice theory is that it mistakes concern with self-interest for rationality. The complaint is well grounded but pointless. Rational choice theorists commonly do suppose that the

* An earlier version of this paper was presented to the political theory seminar at Yale University. I wish to thank Stephen Elkin, Kristen Monroe, James Scott, Ian Shapiro, and participants in that seminar for their comments. I also thank the Mellon and Russell Sage Foundations for generous general support.
agent's own welfare is the chief value of concern to the agent. But there is no mistake involved in such an assumption. It is rather a hypothesis than a direct finding. If it produces good predictions and explanations, it is a good hypothesis; otherwise it is a bad hypothesis. In any case, it cannot be judged a mistake merely by direct inspection of motivations. The critic should want to show not that the assumption is wrong by inspection but rather that it leads to wrong predictions and explanations.

It may be that such a theory will fail too often to be of explanatory value. In actual fact, such a theory in the hands of Bernard Mandeville, Anthony Downs, Gary Becker, and Mancur Olson seems to have swept still less successful theories from the field in various contexts. The vast literature showing that Olson gets interest groups wrong is a literature about the success of tiny fractions of people in various potential interest groups who have taken some degree of collective action. A theory that fits the overwhelming bulk of people most of the time and only fails for small fractions is so unlike any other theory in the dismal history of the social sciences that, if it were Mendel and his theory of inheritance in peas, we would think the data were phoney. That they are not phoney is best attested by the work of the critics themselves who show very low levels of contribution toward various collective provisions.

Oddly, the critics often seem to want to have their cake and eat it too. They criticize moves by many rational choice theorists to include broader concerns in individual utility functions. For example, William Riker and Peter Ordeshook include a sense of "citizen duty" in their list of "rational" motivations to vote (Riker and Ordeshook 1962) and Gary Becker (1962) says the utility function can be quite catholic in its inclusions. Critics say such moves reduce rational choice theory to tautology, making it a matter of constant redefinition to fit what happens rather than a matter of prediction or explanation. But they also complain that a theory limited to self-interest is far too narrow to explain individual behavior and social outcomes.

In any case, the power of rational choice theory has come from its clear assumption, most of the time, of a substantive value that drives individual choices. That value is often captured in the rubric "self-interest," although that is the wrong term for a value. I have an interest in having lots of money. Money is only an instrumental, not a consumption or welfare value. It is a means, not an end. Most of my interests in

---


contemporary social science contexts are instrumental means, as money, power, and efficacy in voting are. The underlying value of rational choice theory is generally own welfare. This value theory gives rational choice its bite. Without a substantive value theory, it is only conceptual, not explanatory.

A quick survey of contemporary rational choice theory

There are two main classes of results in rational choice theory, one of which is primarily conceptual, the other of which is primarily explanatory. First, Kenneth Arrow's impossibility theorem says, in brief, that under well-specified conditions a group's interests cannot even be defined in the same terms as those that define an individual's interests. It is a fallacy of composition to speak of collective or group rationality as an analog of individual rationality.

For the second result, Downs's economic theory of democracy explained why we might get tweedle-dum and tweedle-dee candidates in competition, why voters might not know enough to vote their interests, and why many might not vote at all. Olson, in a variant of the last of Downs's points, showed why interest groups may not form at all to represent many, especially widely popular, interests and why the political interest group system has massive biases in favor of groups already organized for other purposes. Hence, it is also a fallacy of composition to suppose a group can act or be motivated to act as an individual can.

Note that Arrow's result requires no specific value theory. It follows for virtually any value theory that is a preference theory, i.e., that orders states of affairs. Downs's and Olson's results carry for any case in which we might say there is a collective benefit whose provision would benefit all the individuals in a relevant group, in which the costs of providing the collective benefit must be borne by the individuals, and in which the costs and benefits can be aggregated into net benefits. Hence, these results apply to cases in which there is a relatively clear substantive value at stake. In what follows, the concern will generally be with explanatory rational choice theory, which has generally been built on a relatively well-defined value theory.

In the sections below, I will canvass contemporary rational choice theory and then discuss a major confusion in what is at issue. I will then discuss proposals for a rational choice theory without substantive values and three non-welfarist value theories that could alternatively underlie social choice and explanation. Finally, for the problem of group, indeed,
revolutionary cooperation, I will consider the chief candidate theory of politics that is not methodologically individualist: structuralist political economy. I will conclude with brief remarks on the explanatory power of the assumption of individualist welfarism.

The value theory of rational choice theory

One of the striking features of the individualist school of political economy is that it starts from a value theory: individual welfare somehow conceived, including, at its most nearly vacuous, simple preferences. This value theory may be objectionable on various grounds, sometimes including considerations of its internal consistency, but it is relatively coherent and it seems to have strong and coherent implications. When it is coupled with a modest variant of Hobbes's or any sociobiologist's psychological theory - that the first or strongest interest of the individual is the self - it leads to a remarkable blending of positive and normative conclusions. Indeed, this was the great invention of Hobbes. By putting his positive theory of human psychology together with his value theory of human welfare, he was able to devise a positive theory of government that was at once seemingly a normative theory. That is a trick many economists seem to think they turn every day.

Hobbes and then Locke insisted that politics centrally address welfare issues. Hobbes thought religious differences should be subordinated to political order, that individual citizens should, in their own interest, let their sovereign decide their religious affiliation. Locke proposed leaving religion out of politics in order to enable politics to focus on issues of property, including property in oneself. Hobbes and Locke have probably been the two most influential political philosphers in the modern western tradition. As with the best of the rational choice theorists, a chief reason for their influence and the intellectual coherence of their arguments is that they start from a particular value theory.

It may seem odd to modern positivists to say that an explanatory theory gains coherence from a value theory. But that is genuinely the story with individualist political economy. This makes sense if it is true that people are motivated by values in their choices, especially if they are motivated systematically by some set of values. Implicitly, at least, most economists and most rational choice theorists are in the tradition of Hobbes and Locke in assuming that people are so motivated. If their assumption is largely true, then rational choice theory should yield good explanatory results. Those who criticize the particular morality of the theory on the claim that people are more driven by other values should not find it objectionable per se to ground explanatory theory in value commitments.

What difference has economic reasoning actually made in political science? Obviously, it has contributed to major revisions in the areas that Arrow, Downs, and Olson addressed, and more recently in the study of congressional organization, regulation, and campaign contributions. In addition, I wish to assert what may be a controversial claim: Economic reasoning has pushed us toward more closely relating normative and positive concerns. For example, it has pushed us into giving normative accounts of institutional arrangements. We assess the institutions for how well they fit our purposes. Earlier generations of social theorists made normative judgments of institutions, but they often drew their normative principles from elsewhere, not from the panoply of positive principles they used. As in Hobbes, so in current work in individualist political economy, the positive and the normative are driven by related motors: individual incentives for individual benefits. This core concern is, moreover, a purposeful organizer of manifold problems and analyses across a broad front. Theory grounded in it may often falter and give seemingly wrong results, but it seldom fails to yield some kind of result.

The argument here is not one of a necessary or conceptual relation between the value theory assumptions and the rational choice structure of the explanatory theory. It is, rather, a causal or explanatory claim. Practitioners of the individualist approach most often assume, with Hobbes, some version of self-interest as the goal the individual seeks to serve. Downs and Olson produce their striking results from this assumption.

But the assumption of self-interest is not necessary to the individualist approach. For example, it does not lie behind Arrow's impossibility theorem, which is arguably devastating for democratic theory. Arrow's result does not depend on self-interest and could still apply even if everyone chose from purely public-spirited motivations. Imposing reasonable structures of self-interest to limit the range of plausible choices actually helps to avoid Arrow's impossibility result in specific contexts. Indeed, even a moral theorist concerned with action from non-self-interested motives may wish to start from a framework of rational choice in understanding individual moral choices. The moral theorist may even

3 Some may consider that a fault, not a virtue in the approach.
wish to juxtapose self-interest against other motivations. And that theorist may often be forced to conclude that self-interest is too powerful in that contest – it gives the better explanation of behavior.

This last observation suggests what is perhaps the greatest source of controversy over the explanatory and moral propriety of individualist political economy. It is not only the individualist focus, it is also the nearly universal assumption of self-interest as the major explainer of behavior that rankles. We seem to know from introspection that self-interest does not rule our lives so extensively as it rules in individualist political economy. Some academics can look within themselves and seldom find self-interest at work. Rational choice theorists seem to take special pleasure in showing, or at least in making plausibility arguments, that much of what we thought was moral or otherwise not self-interested behavior is arguably self-interested after all. Sometimes the answer to this debate is that both sides are right but in different ways. The self-interest theorists may capture much of what people do politically even in extreme contexts and the overwhelming bulk of what they do in more ordinary contexts. But individuals acting from other commitments spearhead many important activities, such as revolutionary, charitable, and ideal-regarding actions. Therefore their commitments are a very important part of explaining what happens even if they are statistically uncommon commitments.

Interests, consumptions, welfare

It is common to say that people are rational in some contexts and not in others. One might be relatively rational in considering mortgages for a home but not in considering partners for marriage. It is merely an extension of this view to suppose that in some general contexts, such as different cultures, people may be not rational even though people in other cultures might be. For example, James Scott has argued that the very conservative peasants of southeast Asia are not rational. Rather, they are profoundly risk averse and their choices are governed by a "moral economy." What he means is that they are not rational in the sense of simply maximizing their production of rice. Instead of adopting seed grains that would have much higher average annual yields, they adopt very hardy, stable grains that will almost always produce at least enough that they will not starve.

When Scott’s thesis is summarized, it sounds like a rational choice account. Peasants are not interested in maximizing resources only in the sense of producing the most total food over the years or the most income from it. They want such resources, after all, only as means, not as an end. The end they want is a reasonably good life. The goodness of life depends not only on resources but also on consumptions that the resources make possible. Life with reasonable levels of consumption seems eminently preferable to a life of fluctuation between great plenty and starvation. Hence, the rational peasant might well choose Scott’s risk averse policy.

Scott supposes peasants are smart but not rational in their risk aversion. One might sooner conclude that rational choice fits the conditions of scarcity, such as those of Scott’s southeast Asian peasants, particularly well. It is in open-ended contexts in which available resources permit fulfillment of a wide range of values that rational choice might seem dubious. Or, at least, rational choice grounded in the narrow value of own welfare might seem dubious. I will return to this point below.

Scott’s analysis is essentially Hobbesian in the following sense. Scott supposes peasants will first make sure of survival and only then go after greater wealth. Hobbes supposed we must all put order uppermost and therefore we accept government. In both cases, there is no tradeoff with the rationality of amassing greater wealth; there is merely the prior need for the sine qua non of survival, without which there can be no benefit from wealth. If the conditions faced by Scott’s peasants are as he describes them, their risk aversion is not a matter of "moral" economy but only of rational choice.

Scott speaks of "preferences which do not make sense in terms of income alone." This is nothing unusual – preferences are coherently seen only over whole states of affairs, in which income is only part of what is at issue. I do not have a preference for blue. I may have a

---


6 I will not discuss the whole of Scott’s theory of his peasants, and one might suppose he has claims for putting other behaviors under the rubric moral economy. I think that all his important arguments, such as mutual insurance and reciprocity, are readily framed as directly rational, as is seemingly extreme risk aversion in the choice of seed grains as discussed here. The latter is, in any case, the most compelling part of Scott’s empirical analysis.


8 Scott, Moral Economy, p. 35.
preference for a blue shirt with certain other clothes to be worn in particular circumstances. As is presumably true of Scott and virtually everyone else as well, I similarly do not have an unrestricted preference for higher income. If higher income entails giving up my academic life or entails giving up more of my leisure time, I may not prefer it to my present income with my present lifestyle and consumption pattern. There are no preferences that “make sense in terms of income alone.”

The issue here is a conceptually messy one at the very foundations of rational choice theory in the welfarist school. We speak, sometimes almost interchangeably, about interests, consumptions, and welfare. But these are conceptually quite different. Our interests are what put us in a position to consume and consumption typically brings welfare. I have an interest in amassing resources, but resources are of no value per se – I want them only in order to be able to consume. Obviously, interests and consumptions trade off with each other. If I consume some things, I must expend some of my resources. (Some social theorists argue that we should ground our normative theories in resources rather than in welfare, that certain conceptual problems in welfare make it finally an unworkable normative principle. Without resolving that issue, we may all readily grant that resources are means without intrinsic value and that what gives them instrumental value is the welfare they can bring us.)

When Scott speaks of a preference for income alone, he misuses the concept of preference. I can prefer one thing to another all else equal or I can prefer one completely defined state of affairs to another. I cannot simply prefer higher income – I must add what I prefer it to and I must add some rough equivalent of the phrase “all else equal.” Similarly, Scott’s peasants can prefer the whole state of affairs that follows from planting a highly productive but also highly fragile strain of rice to the state that follows from planting a safer but less productive strain, or vice versa. If I have either of these preferences and I act on it, then I, Scott’s peasant, am rational.

The beauty of Scott’s account is that, after we correct the conceptual confusions in it, the empirical case still stands its ground. Scott’s analysis translates into a rich rational choice theory of peasant behavior. The principal errors are evidently not in the empirical account but only in the interpretation of their economic significance.

---


Incidentally, “moral” in this account of moral economy seems to mean no more than that, given their established practice, which is eminently in their interest, the peasants attach moral vocabulary to the practice. As do people in many other contexts as well, they deduce an “ought” from an “is.” Or perhaps they merely short-cut their reasoning and moralize means to the ends they value.

Finally, recall the point made by some critics (especially philosophers) of rational choice that we should expect rational choice theory to fail when resources become profligate enough to allow the fulfillment of ever more diverse consumptions. At that level of resources, resources or their amassment may compete with consumption (they typically compete in the sense of trading off resources for consumptions). I may be able to gain more resources only by forgoing consumptions. Moreover, my consumptions may become much more, not less competitive, just because my indulgence of them is less constrained. Under the grim constraints of subsistence, I cannot really choose much beyond limited consumption with survival and more spectacular consumption briefly on the way to not surviving. It is not that rational choice theory becomes increasingly unstable with increasing resources, but that preference over the vast range of possibilities becomes more idiosyncratic and perhaps even less stable.

---

Rational choice without substantive values

There are several ways in which theorists avoid the assumption of welfarism (or any other particular substantive value) as the basis of rational choice. Two are especially instructive. In the first, it is sometimes assumed that each individual can have any values, welfarist or otherwise, and that the theory will still work, as though the theory were somehow neutral with respect to values. In the second, it is supposed that we get

---


13 In his discussion of “the moral economy of the English crowd,” E. P. Thompson’s central concern is with the reaction of the general populace to high prices for grain in years of short harvests. In those years, the people insist, often under threat of force, that dealers and millers not charge substantially higher prices than they would in good years. This is a “moral” concern insofar as it is an objection to profiteering from adversity. E. P. Thompson, “The Moral Economy of the English Crowd in the Eighteenth Century,” Past and Present 50 (1971): 76–136. Thompson sometimes frames his argument as a derivation of an ought from an is – as did, no doubt, the eighteenth-century English crowd. The price in years of normal harvest becomes the morally right price.
stable results without any assumption whatever about the internal mental considerations of the actors as though the theory were purely behavioral and not implicitly cognitive.

First, consider the neutral view of values. Arrow's theorem is based on leaving it open to each individual to have any preference ordering over various states of affairs. My preferences could be ordered by an altruistic, selfish, spiteful, ideal-regarding, or any other principle. This, the most abstract form of the utility function — with any values plugged in — is a late, essentially theoretical development. Actual economic reasoning about society typically continues the traditional presumption of welfare values, the presumption that there are objective things that matter to most people. In actual reasoning, for example, it is typically assumed that people prefer more resources to fewer, more of some consumptions within a given range to fewer, reasonable survival to death, and so forth. Many applications of economic reasoning would make no sense without such assumptions, which are so much a part of second nature that they are not even explicitly stated.

Similarly, claims for the inherent neutrality of the liberal program are prima facie false or at least exaggerated. Virtually all liberals and all liberalisms share Hobbes's view of the substantive values of survival and welfare. Hence, there is real, substantive content to the values defended by liberalism, which is not an empty shell that can be filled with any values willy-nilly. Indeed, Locke, an early liberal, was particularly concerned to keep non-welfarist religious values out of politics in order to secure welfarist values. The only neutrality we might insist on is the neutrality of John Stuart Mill's skepticism in the determination of what conduce to the welfare of other individuals. One person thinks parlor games are the best way to enhance welfare at the end of the day, another thinks relaxing mindlessly before the television is the best, another thinks reading great classics is best. All may be right for themselves alone. The fact that you think television is best for you is prima facie evidence that it is best for you. I must be epistemologically neutral about which is better for you in principle and must therefore defer prima facie to your judgment unless I can find causal relations that might sway your views.

Millian skeptical neutrality is not neutrality with respect to welfare, virtue, or whatever choice of fundamental value. It is epistemological, not moral. It is neutrality with respect to judgments of what contributes to welfare for a particular individual. Some things can be shown to have systematic welfare tendencies, as health does or various resources and talents might. But any such claim for systematic tendencies is empirical, not a priori, and it can prove to be wrong in a particular momentary

instance or even more generally for a particular person. Hence, these claims may still be overridden by other considerations, such as particularly grim conditions that might make life seem worse than death, or harmful causal effects of which the particular individual may be ignorant. Enjoying wine this moment may seem to me to be a good thing, but it may contribute to wrecking my life.

Although the value theory of most rational choice theory is not neutral, the theory itself is neutral with respect to substantive social problems it can address. It can address almost any problem, almost every kind of problem. It is not a theory of class conflict, state formation, revolution, or democracy. But it can be brought to bear on all of these.

Now turn to the purely behavioral view of choice — choice that is no longer rational in the usual sense, because it involves no mental suppositions about agents. Gary Becker has argued that we do not need the assumption of rational choice to get standard results in economic analysis. He analyzes the demand curve on the assumption that resources are limited rather than on the usual assumption that choice is deliberately intended to maximize utility. He shows that the market demand curve is negatively sloped even if individual households are irrational in their consumption choices. Hence, Becker demonstrates that we can get market regularity from individual irregularity.

There are three, perhaps too easy, responses to this move. First, Becker himself does not follow up his claim thereafter but sensibly returns to the more common assumption of rationality. He defends this move on the compelling ground that the standard theory is so richly developed that the gains from working with it in his areas of concern far outweigh what gains might be had from attempting to create a new theory.

Second, rationality that is inherently grounded in substantive values of welfare is closely related to resourcist accountings. The appeal of basing theory in resources is that, ostensibly, the theory is then fully behavioral. We need not posit mental principles for choosing. We merely lay out the resource constraints. But this apparent escape from the supposedly hidden world of the mental requires at its outset that we somehow intuit or know how the resource constraints matter. We start from the mental, we do not escape it. Becker's demonstration was not directed at the philosopher's constant concern with the mental and the behavioral but only at the understanding of large systemic patterns.

Rationality and *homo economicus*

Third, in politics it is often not sensible to characterize collective choices as choices by individuals in a market. The choices are very often of collective provisions, not individual provisions. We might still expect resource constraints to affect our collective survival, just as individual resource constraints affect individual survival. But this will be a big general issue, not an issue relevant to every collective choice. If we wish to understand many particular collective choices or if we find the results of collective choice relatively stable or regular, we are likely to be driven back to assuming the individuals involved in the choices are acting rationally, not merely whimsically or incomprehensibly.

To a large extent, unless we can carry out a program such as Becker’s, the difference between assuming some value theory for the actors and not assuming any is the difference between explanatory and conceptual analysis in individualist political economy. If we impute certain substantive values to the actors in our theory, we can say what behaviors or choices follow from those value commitments. If we do not, we cannot say much. Hence, we may treat rational choice theory as a strictly positive theory in the following sense. The theorist takes the values of individuals into account but the theorist need not take a value position. If a theorist seems strongly committed to the descriptive claim that all or most people share some value, one might suspect the theorist shares the value. And some rational choice theorists may seem to hold the view that self-interest is the “right” value for a person to have. But one could be completely open as to whether there are any right values while still thinking, descriptively, that some small set of values motivates a particular population in some context.

Non-welfarist value theories

Reliance on a substantive value theory is, of course, not unique to rational choice theory. Theories, both normative and explanatory, have been grounded in various non-welfarist value theories with specific substantive content. Canvassing many of these or analyzing any one of them in depth would take up far too much space for present purposes, but it is useful to consider three major traditions: contractarianism, Marxist labor theory of value, and communitarianism. All of these have been applied both to normative justification and to explanation of social results.

The normative core of rational choice theory

Contractarianism

In its explanatory mode, political contractarianism was ridiculed by Hume and it has never recovered. Hume merely pointed out what everyone knew: that no actual states seemed to have historical grounding in an act of general popular agreement or contract. Still, contractarianism, the political generalization of consent, is touted as a justificatory principle.16 In this, it trades on the appeal of consent in ordinary personal contexts. But that appeal does not compose readily into a collective analog. The apparent moral force of contractarianism is a fallacy of composition. There might be a context in which composition from individual to collective consent would work. But modern states are not such contexts.

Hobbes clearly wanted justification. Perhaps therefore many interpreters have tried to impose a moral contractarian vision on him, as though he were arguing that because we have agreed to government we are morally obligated to stick by our agreement. His actual justification is more nearly utilitarian because it is grounded in mutual advantage. He says we are all better off to have a state and, once we have one, to avoid dissension and revolution. Mutual advantage is, of course, part of what is required for a standard contractarian argument to go through. But Hobbes did not need to take the further step of claiming that prior agreement makes obedience morally obligatory because mutual advantage was sufficient justification for his argument, as it often would be for a utilitarian justification.

Labor theory of value

Turn to the labor theory of value. The chief candidate for Marx’s value theory is his labor theory of value, according to which the value of an object is a function only of the amount of labor that has gone into producing it. That value theory is conspicuously silly in general. For example, if academic journals applied that theory to the acceptance of articles for publication, the journals would deteriorate dramatically. But if that value theory fails, the Marxist argument about exploitation collapses with it. One might still think there is capitalist exploitation of

Rationality and *homo economicus*

workers or first world exploitation of the third world, *but it will take a new argument grounded in a different value theory to make the case.*

Is the labor theory of value strictly normative or is it also explanatory? It would be out of character for Marx not to have treated it as explanatory. The account goes roughly as follows. Workers in some circumstances, such as the nineteenth-century industrial factory, begin to recognize their actual state through discussion of common interests. In that state, they are exploited because they get only a fraction of the labor value they contribute to the capitalist’s product. As their understanding grows, they achieve class consciousness, they fully recognize their common interest as a class in opposition to another class. They are therefore then in a position to act in common for their class interest when relevant opportunities occur.

If we assert there is exploitation without grounding the claim in a value theory, such as the labor theory of value, the claim is likely to involve a naturalistic inference to a normative conclusion. We simply look at the structure of payoffs in an interaction and infer exploitation from nothing more than these facts. That would be an illegitimate and unconvincing move. Therefore, the labor theory of value or some alternative is necessary for Marxist argument, whether normative or explanatory. That is why, if the labor theory of value proves to be incoherent, then Marxist exploitation is incoherent as well – unless, again, it can be given grounding in some other value theory or normative principle.

Incidentally, it should be clear that one can reject the labor theory of value and claims of exploitation based on it and still argue for greater equality or for redistribution of resources for welfare. Indeed, exploitation *per se* is an odd notion that seems inherently to involve a naturalistic inference of a normative claim from a factual assessment. We might readily conclude that there is no exploitation but that there is grievous inequality or deprivation that merits correction. As the demand for unskilled labor falls in the United States, the permanently unemployed underclass grows. They cannot sensibly be seen as a capitalist resource, a reserve army of the proletariat available for controlling wages in the most profitable industries, although they may have that function in some service industries such as some fast-food restaurants. It is increasingly implausible to view their plight as one of exploitation. Most capitalists might be better off if there were no underclass.

Oddly, Marx’s own value theory seems likely to be regressive, not progressive. For example, labor values cannot typically be expected to trump exchange or market values in a free-wheeling market. Suppose therefore that we leave it to central agencies to assign labor values in order to secure relevant incomes based in these values. These are likely to be grounded in past ways of doing things. Hence, in a period of technological transition (which is virtually always), we will tend to assign values for the manufacture of, say, cloth that derive from earlier values when there were no modern, mechanized looms. We will be Luddite. If we are, then we perversely undercut the possibilities of economic welfare for the large masses, whose poverty seemed to drive Marx’s theorizing.

Communitarianism

Finally, consider an older view of what we might call cultural motivation, a view that is now articulated in communitarian thought. Communitarian argument is predominantly normative rather than explanatory. Historically, cultural motivations have been used in explaining collective action by diverse ethnic, religious, local, and national communities. In a sense, the appeal to communitarian values allows us to escape the fallacy of composition inherent in generalizations from individual to collective interest by supposing that there is no problem of composition of individual values.

Superficially, there seems to be no fallacy because every individual is supposed to want the same thing and to have the desire to contribute to achieving it. On the contribution side of the claim, however, one may still expect substantial conflict, just as in ordinary collective action contexts in which we all want the same collective provision but it is in our interest individually that others do the providing. Hence, communitarian values do not resolve the problem of composition from individual motivation to collective action. Their role is rather to admit non-welfarist values into the individual’s panoply, values that are determined somehow by the group.

Some community values are merely group-level aggregations of individual-level values. For example, individuals may be expected to benefit from general economic growth in some contexts. Such welfarist values seem unlikely to be problematic for rational choice theory unless they are reified at the individual level in the form of strictly group-regarding values. Some community values, however, need not be welfarist – although it is too facile to conclude that they necessarily are not welfarist merely on the testimony of those who hold them. For example, many people may assert that the only reason to keep a promise is that it is immoral on non-welfarist grounds to break one. And yet they may generally only make promises for welfarist reasons and may typically break them for welfarist reasons.
Much of recent work on communitarian values is justificatory. The purpose of the work is to argue that communitarian values are good in some sense and that they should be honored by individuals and polities. Those who are troubled by news of atrocities and bloodshed in the name of particular community values have solid ground for complaint against too quick communitarian justification. Much of the work has been conceptual in trying to make sense of how values can get constituted at all if there is no communal basis for at least some of them. So far, the “theory” is woefully underdeveloped. The relatively casual blending of Aristotelian virtues with Hegelian community commitments that stands in for communitarian value theory is too bland and unstructured to yield systematic conclusions. Worse still, the value theories of contemporary communitarians seem to share the regressive bias of the labor theory of value. Those theories elevate a form of social Luddism to the level of the good for a given community.

The structuralist alternative to individual choice

Explanations grounded in these alternative value theories are still methodologically individualist. In the face of major political phenomena such as revolutions, the structuralist political economist may assert that an individualist approach is inadequate even on the accounts of individualist political economy. If the bulk of behavior would not be revolutionary and only a relatively small number of individuals make a revolution go, we cannot expect to know enough about relevant individuals to make competent predictions or even retrospective assessments of revolutions. We cannot have a theory of revolution or great social movements that depends on predicting whether a specific Lenin or Gorbachev is available to lead, and we cannot believe Leo Tolstoy’s functionalist vision, at the end of War and Peace, that great times necessarily produce great people. We will have to look to more systematic structural considerations that finally enable the occasional revolutionary to make a difference.

One answer that the individualist theorist can give to this criticism is that the revolutions we see may be more or less accidental rather than structurally determined. Hence, the structuralist approach does no better than the individualist. Structuralists are often uninterested in individual commitments and are willing simply to assume that relevant kinds will manifest themselves under certain conditions. They can then focus on explaining or understanding those conditions. The great strength of the individualist school is in simple generalizations about individuals, generalizations that may transform understanding even while blurring differences that seem to matter historically.

In the recent gentle revolutions of East Europe we may see seeming vindication of both individualist and structuralist arguments. For many commentators, the structural conditions of apparently declining economies seem to explain the revolutions. The accident of Gorbachev seems to many others to be the chief explanation of at least the timing of the revolutions. But individualists might further argue, as most western economists do, that the structural conditions of declining economies are eminently easily explained from individualist assumptions. The structuralist approach is therefore, at least in large part, merely shorthand for a more precise and more extensive individualist account. Moreover, individualist economists were perhaps more consistent than any other social scientists in predicting the changing conditions of the political economies of the East European nations. To trump them, structuralists would have to argue that the major condition was not economic but rather political and moral — it was the urge for democracy and openness, for national self-determination and the end of colonial domination, or whatever. Ironically, that would be an odder move in the structuralist school than in the individualist school.

Concluding remarks

If theorists of other persuasions are to take on individualist theorists on an equal footing, they too will have to accommodate their theories to a compelling value theory, which at present levels of theoretical development of value theories, means either to accommodate to something like the individualist value theories of welfare and preference or to develop some other theory far more extensively than has yet been done. If the theory remains hollow at its core, it will be only variously of interest in its occasional insights rather than generally of interest in its overall program.

Perhaps rational choice theory has a natural tendency to develop a value theory roughly like that which drives most of it. Because it is individualist, it must be concerned with individual actions, hence, most
likely, individual motivations. What motivates an individual? Typically, the desire to achieve particular ends or values, perhaps especially self-regarding ends and values, as rational choice theorists generally assume.

Rational choice theory is finally like sociobiology in the following limited sense. The most obvious sociobiological inference about human motivation is that humans seek their own interests. If they did not, they would individually reduce their chances of survival. This inference leaves open for investigation whether it could happen that group-oriented behaviors get selected. Economic reasoning about human motivation starts from the simplest self-interest assumption and then turns to the same range of more complex aggregate or group-level issues. Still, in both cases, the group-level results are derived from individual interests. Even at the group-level, we would expect the individual-level result still to hold — if not always, at least very often, perhaps even predominantly often.

Much of the power of the rational choice approach may be an accident. It just happens that the welfarist value theory of economics over the past two or three centuries grew in response to theoretical and empirical attacks from vast numbers of economists and their critics. Nothing even vaguely like that has happened to any other value theory. As a result, the formal value theory of rational choice is rather articulated and supple. Anyone who wishes to oppose some other value theory to it is burdened with the virtually impossible task of first making that other theory articulate enough even to stay in debate. Worse still, at this stage in development of various theories, all others are virtually forced to deal with the formal and even the substantive value theory of rational choice if they are to address similar problems. That is to say, other theorists are likely to be forced to contribute to rational choice theory while pursuing their own preferred alternative. The critics of rational choice may finally wish to weep at their complicity in its refinement.

19 This was Hobbes’s assumption long before it was a sociobiological inference.

5 The virtual reality of homo economicus

PHILIP PETTIT

The economic explanation of individual behaviour, even behaviour outside the traditional province of the market, projects a distinctively economic image on the minds of the agents involved. It suggests that in regard to motivation and rationality, they conform to the profile of homo economicus. But this suggestion, by many lights, flies in the face of common sense; it conflicts with our ordinary assumptions about how we each feel and think in most situations, certainly most non-market situations, and about how that feeling and thought manifest themselves in action. What, then, to conclude? That common sense is deeply in error on these matters? That, on the contrary, economics is in error — at least about non-market behaviour — and common sense sound? Or that some form of reconciliation is available between the two perspectives? This paper is an attempt to defend a conciliatory position.

The paper is in five sections. In the first section I describe the economic mind that is projected in economic explanation, whether explanation of market or non-market behaviour. In the second section I argue that this is not the mind that people manifest in most social settings and, in particular, that it is not the mind that common sense articulates. In the third section I show that nevertheless the economic mind may have a guaranteed place in or around the springs of human action; it may have a virtual presence in the generation of action, even action on which it does not actually impact. In the fourth section I show that where the economic mind has such a virtual presence, that is enough to license an important variety of economic explanation: the explanation of the resilience or robustness of certain patterns rather than the explanation of their emergence or continuance. And then in a short fifth section I show that this sort of explanation fits with some established ideas about the explananda of economic and social science.