

REVIEWS

doi:10.1017/S0266267105210672

What Price the Moral High Ground? Ethical dilemmas in competitive environments, by Robert H. Frank. Princeton and Oxford: Princeton University Press, 2004, xii + 203 pages.

What price the moral high ground? In abandoning the methods of science for those of journalism, Bob Frank pays a price that is too high for me. I share his distaste for the dishonesty of the rhetoric of right-wing economists, but the answer is not to fight fire with fire in the style of Michael Moore.

It is certainly entertaining to suggest that neoclassical economists are evil, money-grubbing misfits, but mostly they aren't. Nor is homo economicus someone who selfishly cares only about money. Nor do evolutionary considerations somehow trump neoclassical reasoning. On the contrary, it is partly because the methods of neoclassical economics have been imported into evolutionary biology that theoretical biology has taken off so spectacularly in recent years. It is true that human nature is much more complex than the likes of Milton Friedman are willing to admit, but to use him as a spokesman for contemporary neoclassical economists is grotesque. Many of us spend much of our time using evolutionary game theory trying to make sense of the very quirks of human nature that Bob Frank implies that we ignore.

What Price the Moral High Ground? is a collection of essays that have been widely influential in discrediting modern economic theory among those in the intellectual community at large who “always knew” that something was rotten in the State of Denmark. They are written with great panache and style, skillfully using anecdotes and carefully chosen statistics to press home points that pander to the prejudices of the general reader. As pieces of literature it is hard to see how they could be bettered.

The first essay defends the Transparent Disposition Fallacy, which says that human beings are able to commit themselves to actions in the future

that they will not wish to carry out should the occasion arise, *and* that they can convince other people that such a commitment has been made. David Gauthier's (1986) appeal to the fallacy in his risible attempt to show that cooperation can be rational in the one-shot Prisoners' Dilemma (Binmore 1994) is mentioned with approval.

Frank follows up an idea considered earlier by Darwin (1965) in his *Expression of the Human Emotions* that the supposed transparency of our commitments can be traced to muscles in the face over which we have little or no voluntary control. But Frank chooses not to tell us that Darwin actually rejects the idea on the grounds that people can learn to control their expressions by internally simulating the emotional states to which they correspond – as with actors and confidence tricksters. Nor does Frank's (1988) *Passions within Reason* point out that the photograph he reproduces from Darwin's book is the only photograph that isn't posed by Victorian actor convincingly simulating various emotional states.

It is also in this essay that he explains why he disregards what neoclassical economists write in their textbooks about homo economicus. We apparently wouldn't be able to predict anything if we really only required of homo economicus that his decisions are consistent with each other. We are therefore supposedly stuck with the assumption that homo economicus selfishly maximizes the money in his pocket. Perhaps Frank should read the works of behavioral economists like Ernst Fehr (1997) who take the neoclassical paradigm seriously when they fit "other-regarding" utility functions to their experimental data; or economic theorists like Larry Samuelson (2004) who offer plausible neoclassical models of how and why evolution might have made us care about the welfare of others.

The second chapter describes a Cornell experiment in which most subjects succeeded in cooperating in the one-shot Prisoners' Dilemma after they were allowed to discuss how they would play the game with the other player in a half-hour preplay session. This isn't very surprising, since the reward for metaphorically stabbing your partner in the back was only one dollar. But what I want to comment on is Frank's failure to note that his experiment is only one of a vast number of experiments on the private provision of public goods. These were surveyed very professionally by Ledyard (1995) for the authoritative *Handbook of Experimental Economics*. Ledyard's survey confirms that inexperienced and underpaid subjects often do begin by cooperating in games like the one-shot Prisoners' Dilemma, but that the rate of defection climbs inexorably as the subjects gain experience or the payoffs are raised until about 90 percent of subjects are defecting. This inconvenient empirical fact is also relevant to Frank's ninth and last chapter, in which he reports an experiment that supposedly shows that people trained in economics are more wicked than other folk. It never seems to occur to those who go in for this kind of thing that perhaps

an economic training simply short-circuits the trial-and-error learning that will eventually lead most other people to the same conclusion.

There is a lot else in this book, much of which I think probably hits some important nails on the head. But how selective are the anecdotes and the statistics that are quoted? Are we being told the whole story, or only that part of it which suits the author's purpose? Having seen the data in one domain distorted to support claims I believe to be false, how can I avoid feeling that my scientific judgment is perhaps being subverted in other domains by the brilliance of a rhetorical style designed to exploit my vulnerability to popular prejudice?

In summary, this book is a triumph of polemical journalism. If made into a movie it might easily win the Palme d'Or at the Cannes Film Festival. If you think neoclassical economics deserves a good kicking around, you will enjoy it a lot. But if you think of economics as a science, its lack of objectivity will merely irritate.

Ken Binmore

University of Bristol

REFERENCES

- Binmore, K. 1994. *Playing fair: Game Theory and the Social Contract I*. MIT Press
- Darwin, C. 1965. *The expression of the emotions in man and animals*. University of Chicago Press
- Fehr, E., S. Gächter and G. Kirchsteiger. 1997. Reciprocity as a contract enforcement device: experimental evidence. *Econometrica* 65:833–60
- Frank, F. 1988. *Passions within reason*. Norton
- Gauthier, D. 1986. *Morals by agreement*, Clarendon Press
- Ledyard, J. Public goods: a survey of experimental research. In *Handbook of experimental game theory*, ed. J. Kagel and A. Roth, Princeton University Press
- Samuelson, L. 2004. Information-based relative consumption effects. *Econometrica* 72:93–118

doi:10.1017/S0266267105220679

The Economics of Poverty and Inequality, edited by Frank A. Cowell. Volume I *Inequality*. Edward Elgar, 2003 xxxv + 627 pages.

Suppose you have the task of dividing a cake at your daughter Katie's birthday. You may be inclined to give a relatively big piece to Sam, because he did so well at Blind Man's Buff earlier, to George, because he hit his head while playing the same game, or to Katie herself, because it is her birthday. The possible principles of cake division are multifarious, as are candidate principles for social justice more generally. You may be

motivated by considerations of utility, need, merit, or others still. Whatever your currency of justice, however, a principle of *equal* distribution has long been considered particularly salient. *The Economics of Poverty and Inequality*, edited by Frank Cowell, collects a number of classic articles that demonstrate the ways in which economic analysis has incorporated this concern for equality.

Most of the contributions to the volume have been widely commented on, some can even claim to have triggered an entire new literature. So why write a review of such a collection? The purpose of this short piece will be to evaluate the discussion of equality and inequality in the discipline of economics within the broader context of the philosophical debate on these same issues – I will focus exclusively on Volume I, *Inequality*.¹ By pointing to more or less recent developments in the philosophical literature on (in)equality, I will try to convince you that the economic analysis of (in)equality has remained relatively – perhaps untenably – narrow.

We can distinguish three different approaches towards the concept of (in)equality in economics: the welfarist tradition, which is interested in inequality because it may be harmful to social welfare; a number of “objective” measures of inequality like the variance of incomes or the Gini coefficient; and finally, a literature that takes the intrinsic value of equality in distribution as given, and aims to evaluate the implications of different conceptual ways to capture this intrinsic value. We will look at these categories in turn, focusing on the first and the third; at the same time, we will identify some of the contributions assembled by Cowell as belonging to one or other of these categories.² I will show that both the welfarist tradition and the assignment of intrinsic value to equality have met forceful criticisms in the philosophical literature. Within the discipline of economics, awareness of the anti-welfarist critique is also widespread, and it is therefore surprising to find it all but absent from the present collection. The worry in the second case is more fundamental. When confronted with the increasing level of sophistication that has characterized recent philosophical analysis of (in)equality, the way in which economic analysis assigns intrinsic value to equality appears simplistic. Let me back up these arguments one by one.

1. The welfarist approach to (in)equality has a long pedigree in economics, and it has undergone a number of important changes over the years. The first two authors in the present volume, Pigou (1) and Dalton (2), represent the “old” welfare economics that was steeped in

¹ A brief terminological note: Unless specified otherwise, we will treat the concepts of equality and inequality as two sides of the same coin. Economists tend to use the latter term, philosophers seem to have a preference for the former.

² Cf. the numbers in brackets and italics given throughout the text.

the utilitarian tradition and featured interpersonally comparable and cardinal welfares. Note that, in this framework, an equal distribution of income would indeed be recommended under the unlikely conditions of people having identical, concave utility functions. However, in the era of logical positivism, the assumption of interpersonal comparisons was attacked as epistemologically dubious (cf. for instance Robbins 1932) and subsequently abandoned. In the wake of the positivist revolution, welfare economists found themselves thrown back to the concept of Pareto optimality, a tool that is well known to cut preciously little ice in analysing inequality. Two new concepts emerged: *compensation criteria* like those developed by Kaldor or Hicks that aimed to increase the cutting power of Pareto optimality for economic policy; and the *social welfare function* introduced by Bergson and Samuelson, which aimed to make policy recommendations by aggregating individual value judgements. Terminologically, the fall from grace of interpersonal comparisons was accompanied by the substitution of *preference satisfaction* for talk of individuals' *utilities*.³

Cowell chooses to focus on the heritage of the second of these developments. Most of the articles in Part IV (Inequality: Welfare Approach) and some in Part II (Welfare and Inequality Rankings) are dedicated to the notion of the social welfare function and its conceptual ramifications. Atkinson (17) probably represents the "new" face of welfare economics most accurately. His concept of *Equally Distributed Equivalent Income* preserves the original welfarist intuition that inequality is harmful to social welfare. Inequality, according to Atkinson, should be measured by asking the following question: What portion of our current level of income would it take to obtain the same level of social welfare, provided we distributed that portion equally? The smaller this percentage, we may conclude, the more unequal the distribution under the *status quo*. It is worth re-emphasising that, from Atkinson's perspective, the value of equality is purely instrumental.

Before we turn to the critique of welfarism, it is worth mentioning a final twist in the history of the welfarist tradition. Several economists, frustrated with the loss of explanatory power incurred through the theoretical ban on interpersonal comparisons, ventured to rehabilitate them. Harsanyi (3) provides an early and clear example of this movement, Kolm (16) can also be interpreted in this way. Importantly, it is not the theorist who makes interpersonal comparisons, but they are implicit in the "ethical" – as opposed to "subjective" – set of preferences held by

³ The historical sketch of this paragraph borrows heavily from Kotaro Suzumura's entry on "social welfare functions" in *The New Palgrave Dictionary of Economics*.

every member of society. The task of the theorists consists “merely” in the objective aggregation of these subjective value judgements.

The principal criticism of the welfarist view runs thus: The notions of both utility and preference satisfaction unduly focus our attention on the *level* of people’s achievements, neglecting potential differences in their *freedom* to choose between various such achievements. A welfarist principle of distribution will arbitrarily favour relatively good or bad “utility converters” respectively – depending on whether the principle advocates a maximization or equalization of welfare. Under a principle of welfare maximization, for example, someone with a sunny disposition will benefit. Such a disposition, physical abilities and disabilities in general, but also factors in one’s social environment, can act as systematic constraints on one’s freedom to choose between different achievements.

One of the main advocates of this welfarist critique is, of course, Amartya Sen (cf. for instance his 1992). Instead of utilities or preference satisfaction, he argues in the positive part of his theory, we should focus on *capabilities*. Capabilities comprise factors like health, education, and nutrition, and they enhance our capacity and freedom to use resources for our diverse ends. Let me make three comments at this point. First, Sen’s position has a feature in common with the last current of the welfarist tradition mentioned above. Like Harsanyi or Kolm, he thinks that interpersonal comparisons are both necessary and possible. Yet, instead of making them in the realm of utilities or preference satisfaction, Sen proposes to draw them in terms of the perhaps vague, but more easily comparable, capabilities. Second, note that for Sen, like for the welfarist, the distribution of income only has instrumental value. In his case, income should be distributed so as to equalise capabilities as much as possible. Finally, against the background of all this, it is surprising that the welfarist critique is not represented in Cowell’s volume.⁴

Expressed in terms of our opening example of distributing a cake, the insight of this section can be summarized as follows. A principle that distributes cake – or income – exclusively on the basis of the preference satisfaction of the potential recipients is doubly defective. Not only does it neglect the possibility that there may be considerations other than preference satisfaction that call for a different distribution of income, but, more importantly, the notion of preference satisfaction turns out to be an unsatisfactory measure of social advantage. Not only may you be justified in giving a larger piece of cake to George because he hit his head playing blind man’s buff, or to Alex because she is malnourished, but a distribution according to preference satisfaction may

⁴ Some of Sen’s other work is included. Cf. his article co-authored with Partha Dasgupta and David Starrett (18) as well as three articles in Volume II on poverty.

unfairly favour those who happen to be good at converting cake into utility.

2. A second approach to (in)equality within the discipline of economics is represented by a set of positive measures of inequality. The characterization “positive” distinguishes them from measures that presuppose a particular normative perspective, like for instance a welfarist position or one that attributes intrinsic value to an equal distribution of income.

Cowell dedicates Part III (Inequality Measurement) to these issues, though they also play a role in most of the other parts of the volume. One of the most popular inequality statistics is the Gini coefficient (13), which measures the extent to which a distribution of income among individuals deviates from a perfectly equal distribution. Like most inequality statistics, the Gini coefficient conforms better to some of our intuitions about inequality than to others. One of its weaknesses consists for instance in its insensitivity to the pattern of inequality, i.e., the question to what extent overall inequality is due to inequality among specific subgroups of the population, like the rich. Efforts to analyse the structure of inequality in more detail, or making income inequality measures *decomposable* – cf. Parts V and VI of the present volume – are in part designed to overcome this weakness.

When inequality statistics are used in practice, they cannot maintain their “positive” character, but inevitably have to be combined with a justification for why we care about this particular kind of (in)equality. From a philosophical point of view, it is this justification that interests us. Since the literature on inequality statistics brackets this issue, we will not spend any more time on it here.

3. Our list of approaches towards (in)equality in economics is completed by those theories that assume an equal distribution of income to be of intrinsic value. When comparing two income distributions by using one of the “positive” summary statistics, this assumption provides the normative backdrop against which the distribution exhibiting *less inequality* can be described as *better*.

Tungodden (2003: 5) provides a very accurate characterization of the methodology behind this approach. “First, it is assumed that any ranking [of distributions] should constitute a reflexive and transitive binary betterness relation. Second, economists introduce principles that constrain the betterness relation, and we say that a set of principles justify a betterness relation if it provides a characterisation of the betterness relation, that is, if and only if it is the only betterness relation satisfying all these principles.” These constraints on the betterness relation are typically expressed in terms of axioms.

Paradoxically, it is one of the two philosophical contributions that best illustrates this methodology in Cowell's volume. In order to specify the betterness relation, Vallentyne (7) proposes to complement what he regards as certain uncontroversial conditions⁵ with two new ones: (1) Any benefit (no matter how small) to a person who remains below the mean after the benefit is given takes absolute priority (with respect to equality promotion) over any benefit (no matter how large) to a person above the mean. (2) When only individuals below the mean are affected, and the total number of individuals is constant, benefits should be distributed so as to maximize the total benefit.

Let me make two comments on this kind of approach. First, note that conditions imposed on the betterness relation can incorporate more than just egalitarian concerns. As Tungodden's paper illustrates particularly well, various kinds of combinations of egalitarian concerns with aggregative or prioritarian considerations are possible. Whereas aggregative axioms assign a certain intrinsic value to the sum of incomes, prioritarian axioms weigh the concerns of some particular subgroup – typically the worst-off – more heavily. Meade (20), who proposes to “allot differences in distributional weights to the spendable incomes of various classes of person” (356) could be interpreted as advocating a prioritarian position.

As to the second comment, consider the other explicitly philosophical contribution to the volume by Temkin (5). Putting into context his project of comparing different situations with respect to their inequality, Temkin mentions that he thinks “deserved inequalities are not bad *at all*. Rather, what is objectionable is some being worse off than others *through no fault of their own*.” (67). This statement encapsulates the heart of modern egalitarianism. It is incompatible, so I claim, with assigning intrinsic value to an equal distribution of income.

To argue for this claim, let us rewind for a moment the egalitarian debate in philosophy. Thirty years ago, Nozick characterized welfare economists as defending “current-time slice principles” (1974: 154) of justice, and went on to criticize these for their lack of sensitivity to the history of a distribution. In particular, such an approach disrespects the choices that people make about how to dispose of their resources, and therefore amounts to an unacceptable infringement of individual liberty. Whereas this argument was initially put forward by libertarians, egalitarians soon acknowledged its power and began to incorporate it into their own theories. Thus emerged liberal egalitarianism, represented today by philosophers like Dworkin, Cohen, or van Parijs. As expressed in the

⁵ As for instance Pigou-Dalton, which states that equality is increased by any non-reversing transfer from one person to a worse-off person. “Non-reversing” means that the worse-off person remains relatively worse off despite the transfer.

quote from Temkin, liberal egalitarians hold that inequalities should only be rectified if they are undeserved; that is, people should be compensated for social disadvantage only if they incur it through no fault of their own. Against the background of this conceptual move, assigning intrinsic value to equality of income *per se* is no longer acceptable, but we now have to differentiate according to whether the income in question is traceable to people's choices or to undeserved underlying inequalities.

The economic approach to (in)equality discussed in this section, much like its welfare economic cousin, is still subject to Nozick's charge. The way in which conditions on the betterness relation are formulated – be they egalitarian, utilitarian, or prioritarian in spirit – still restricts our analysis of (in)equality to a single time slice. There is no room for the distinction between deserved and undeserved social disadvantage that is pivotal to contemporary philosophical theories of egalitarianism.

Is there a conceptual remedy for this shortcoming of economic analysis of (in)equality? We lack the space here to discuss this question in any depth, but let me at least point to one potential culprit for this unacceptably ahistorical outlook. As Cowell himself highlights in his introduction, “the basic properties of a welfare [or inequality] ranking or ordering are usually ‘programmed in’ by introducing an *axiom system*” (xiv). One axiom that is common to this entire literature is an anonymity assumption, which requires that our assessment of a distribution be invariant to permutations of the members of the population within the distribution. This implies that you have to be indifferent between distributions $X = (10, 20, 30)$ and $Y = (20, 10, 30)$ among three individuals A, B, and C, where the numbers represent units of income. The motivation behind the anonymity assumption is to avoid “noxious” caring, that is to assign some individuals more moral value than others. Yet, can we really arrive at a meaningful normative assessment of these distributions based on the information available? Suppose we learn that A is a farmer and B a gambler; whereas A's harvest was hit by a storm under scenario X, B's income depends entirely on his luck at the casino. With information of this sort in hand, we are unlikely to remain indifferent between the two distributions. Borrowing an expression from Kolm (29), one is tempted to conclude that, when the anonymity axiom is in place, “the observer does not know what he cares for and does not care for what he knows” (529).

When distributing the pieces of a cake, we often follow a convention of equal shares. The distribution of income is governed by a different set of conventions, most prominently a principle that respects the choices people made in producing the cake. Cases of “pure” distribution, where a principle of equal shares may provide an acceptable rule of thumb, are rare in this context. Given the emphasis economists put on individual liberty and choice in their defense of the market, it is puzzling that their analysis of inequality ignores these features.

In conclusion, we have distinguished three approaches to the issue of (in)equality in economic theory: (1) A welfarist perspective that assigns instrumental value to equality; (2) a set of positive summary statistics that measure inequality of income, and (3) a literature that, in its formulation of conditions on the betterness relation, assigns intrinsic value to equality. Volume I of *The Economics of Poverty and Inequality* includes contributions from all three of these perspectives. I have made two principal comments from a philosophical point of view. First, it is surprising that the anti-welfarist critique has not made it into the volume. If nothing else, this demonstrates how strong a grip the welfarist tradition still exerts on economic theory. Second, and more fundamentally, the economic literature on (in)equality, especially when it takes the form of (3), appears to have not yet digested what is now considered the centerpiece of egalitarian theories in philosophy: Inequalities are only bad when undeserved. A precondition to overcoming this predicament lies in heeding Nozick's advice and abandoning the time-slice perspective. To do so, and to thereby admit individual choice as a determinant of income to their normative picture, I have suggested that economic theorists need to revisit the anonymity assumption.

Peter Dietsch

Université de Montréal

REFERENCES

- Robbins, L. 1932. *An essay on the nature and significance of economic science*, 2nd edn. London: Macmillan, 1935
- Sen, A. 1992. *Inequality reexamined*. Oxford University Press
- Tungodden, B. 2003. The value of equality. *Economics and Philosophy* 19:1–44
- Nozick, R. 1974. *Anarchy, state, and utopia*. Basic Books

doi:10.1017/S0266267105230675

Indeterminacy and Society, by Russell Hardin. Princeton University Press, 2003, xii + 166 pages.

In the biblical story of Abraham and Isaac, God speaks to Abraham: Sacrifice your only son Isaac on a mountain high. Abraham is overcome with fear and trembling. No human interest or human value is served by this act, he concludes. Can it still be rational? Clearly not, Russell Hardin would say, because rationality requires that people choose more rather than less value, and by hypothesis, this choice promotes no value. Not even Soren Kierkegaard disagrees with this, although he suggests that rational judgment does not determine the answer. This latter

idea – rational indeterminacy – is at the center of Russell Hardin’s most recent book, *Indeterminacy and Society*.

If Abraham has problems, what about closer to home, when we walk out the front door and face an everyday world of strategic interaction? Rationality should at least guide us through the social world that we know. But Hardin says that rational indeterminacy is a deep and pervasive feature of these choices too. Indeterminacies are all around us, he argues, and any successful social theory must face up to this reality.

Hardin observes: “[People] manage to get through life most of the time. Our task in explaining their successes and failures is often the task of understanding how they choose to deal with indeterminacies that often swamp reason” (137).

Hardin has a point, but he poses a challenge that the resources of his book do not meet. This review explains and evaluates Hardin’s analysis, and offers an alternative perspective. The remarks to follow are offered against the backdrop of a recommendation: this book will capture the interests of anyone who wishes to think further about the concept of rational indeterminacy and the limits to rationally compelling advice for social organization.

1. Rational indeterminacy and the iterated prisoner’s dilemma

When we need to make decisions, it is nice to think that there are clear answers about what it is rational to do, given the situation. Hardin takes for granted an idea about basic rationality, that we should choose more rather than less value and our rankings of the options should be transitive (16). His book is about understanding the implications of basic rationality for individual and social choice in a stochastic and strategic world. The hope is that basic rationality can provide determinate recommendations about what to do, but Hardin is skeptical.

The structure of Hardin’s book is to begin with the prisoner’s dilemmas to illustrate the meaning of basic rationality, to show the importance and frequency of indeterminacies in our social interactions, and to make clear how poorly our leading theories of rationality can explain what is going on within these troubling interactions. His book then surveys many contexts of strategic interaction and problems of stochastic choice where rational indeterminacy plays a role. Sandwiched in the middle are chapters about Hobbes, Bentham, Coase, Kant, and Rawls, and how they either avoid or deal with these rational indeterminacies. This part of his project is interpretive. Like Will Kymlicka’s (2002) effort to interpret social theories as expressions of equality, Hardin’s goal is to survey and interpret social theories as responses to rational indeterminacy. Hardin also weaves together a positive argument that he develops throughout the

book and then defends in the last chapter. He calls his position “mechanical determinacy” and this will be the subject of discussion below.

To begin, Hardin starts with the strategic interaction of a finitely repeated prisoner’s dilemma and reviews the argument from backward induction that recommends defection in every round. The problem with this always-defect recommendation is that we often do much better for ourselves (in the sense of attaining more rather than less value) by cooperating. Always defect may seem like a determinate solution, but, Hardin tells us, it is contrary to good sense because an alternative that gets people to cooperate at least some of the time makes us better off. This is an odd result. The intuition that drives the recommendation to always defect will drive agents to reject that rational choice theory.

There are several points to emphasize about Hardin’s analysis of prisoner’s dilemmas. First, rational indeterminacy is not causal indeterminacy, a problem about physics, or a problem about missing information. No doubt we are often unsure about how to decide because we do not have all the facts. Hardin is interested in failures of reason or theories of reason that do not trace to missing facts. Iterated prisoner’s dilemmas are such an example. Another example is Condorcet’s proof that a principle of majority choice can give rise to cyclical choices and can therefore be indeterminate as a mechanism for identifying a winner.

Second, he believes that prisoner’s dilemmas are central to social life; this makes rational indeterminacy pervasive. He writes, “That game is a pervasive part of our life because it is essentially the structure of exchange. Any economist’s theory of rationality must be able to handle that game” (5).

Finally, Hardin observes, “[cooperating] is not a determinate answer . . . It is at best savvy, not rational according to the principle of basic rationality. Such savviness is what we should actually hope for in social life” (40). What more do we want? For Hardin the question is: what more can we have? Not much. His analysis of prisoner’s dilemmas illustrates a thematic argument throughout the book, that acknowledging indeterminacy gets us to better outcomes than insisting on imposing determinacy where none exists (3). He discusses the nuclear arms race to illustrate how people can feel their way to better outcomes, through contract by convention, without determinacy. Hardin emphasizes that we need to design theories for a real world with many indeterminacies.

His analysis connects nicely with the work of economist Robert Frank, who also has difficulty deciding whether to describe cooperation in prisoner’s dilemmas as rational or not. Frank (2004) finds empirically that people do cooperate, that this can make sense (depending on population distributions), that it is possible for a population to reach a steady-state of cooperators and defectors, and that all of this can plausibly be understood through evolutionary mechanisms. The experimental

literature in economics supports Hardin's thesis of rational indeterminacy and his pragmatic bent toward recommending cooperation, somehow.

2. Mechanical determinacy: what does it get us?

Hardin wants to know how much, by way of creating determinate social theory, can we squeeze from a single normative assumption, basic rationality. He has us begin with the varying interests of persons, a commitment to basic rationality at the individual level, and the need to aggregate these interests in some way for organizing society. He states that mutual advantage is the collective implication of these ordinalist foundations because it can preserve individual rationality for each and every person. Mutual advantage does not get us very far, however, because almost every policy choice implies trade-offs that will not be to everyone's advantage. His signature example is vaccination policy, which will not advantage those unlucky few who vaccinate and contract the disease, especially given that they may not have contracted the disease if there had been no vaccination policy. The question is: how should we make social choices in circumstances where the choice does not determine a mutually advantageous outcome?

Hardin's short answer is that we have massive indeterminacy and there is no basis to make the choice. "Not much in the way of staking out policy positions can be wrong in political debate," he writes (52). To a large extent this is Hardin's answer, and to a large extent this is the problem with Hardin's analysis: it gets us nothing. The balance of Hardin's project, however, is to suggest that his framework gets us *something*, but here I will argue that we do not even get that.

Hardin endorses the structure of Hobbes's analysis: "Hobbes [went] as far as may be possible in grounding political order in individual interests rather than in an *ad hoc* claim of normative commitments" (54). The analysis is that, *ex ante*, we would each be better off with government than without. Hobbes understood that this does not follow without additional assumptions, which Hardin describes as brilliant tricks for overcoming indeterminacy. We suppose that we have insufficient knowledge to be confident that any particular form of government is better than any other, but we do have sufficient knowledge to be confident that any transformation of our current government is too costly to justify. If so, support for the extant government can be a collective implication of individual rationality.

This is stage one of two-stage social theory. The second stage is to recognize that particular policies will rarely be mutually advantageous, so we institute mechanisms to resolve these conflicts for us. Deploying these mechanisms is *ex ante* mutually advantageous, even if we are unhappy with the results. "Indeed, at the end of a sound social theory stands

an institution," he writes. "In personal contexts we simply live with indeterminacy, and *the passage of time renders personal choices determinate after theory fails* . . . In both personal and social contexts, we therefore often do not achieve theoretical determinacy, but only *mechanical determinacy*" (his italics, 121). Mechanical determinacy is Hardin's own prescription: letting the rules of institutions grind out the results after the government is set up.

An inspiring model? Not really. But Hardin thinks it is the best we can do, and it is *something*. Unfortunately, already, it is too much. Recall that he advocates mutual advantage (7). He does not defend this principle, however, and unfortunately, it is already a substantive commitment beyond basic rationality. Suppose that my own interests are to seek my own advantage. In fact, if I see that it advantages someone else in the process, I recoil in horror and find some other course of action that at least makes someone worse off. I am Pareto's evil twin and I seek Depravo optimality, wishing to make myself better off while making all others worse off. That would be quite an accomplishment, and unfortunately, history teaches us that some people seem willing to try. Pareto depravos reject all opportunities for mutual advantage.

In defense of Hardin, his project is not about a normative defense of his normative commitments (which some readers, including this writer, may find frustrating), but the point is that he cannot even get his ultra-thin results without extremely generous empirical speculations about peoples' preferences. There is no practical mutual advantage argument, even *ex ante*, for anything of real substance that involves millions of people, especially like establishing government. People with wild interests are everywhere; it probably only takes a handful of Pareto depravos to wreck any mutual advantage argument, and, in any case, preferences are just too diverse to get the result he wants.

There is not much place to go if our building blocks are human interests, basic rationality, and massive rational indeterminacy. The mechanical determinacy of institutions is not even pragmatically justified on his framework, given that we have no reason to believe that *any* generalized mutual advantage argument can succeed. Hardin's discussion shows us that if we take rational indeterminacy very seriously and we are committed solely to his ordinal welfarism, we get nothing.

3. Moral anchors about actions

Let us suppose, on the contrary, that Hardin's project gets us what he wants. His best-case scenario is two-stage theorizing with massive rational indeterminacy. He writes, "strategic indeterminacy . . . wrecks any hope for serious agreement beyond, with luck, a few general principles, institutions, and laws and maybe some happenstance policy choices" (51). He says that

“the chief problem of ordinalism has been its indeterminacy, [but] . . . its indeterminacy merely mirrors reality” (70). Mechanical determinacy, he believes, is a good pragmatic response to this reality.

An important aspect of this two-stage approach, he notes, is that “it blocks direct assessments of the rightness, wrongness, justice, or injustice of various states of affairs and actions” (125). In other words, Hardin has a dim view that assessing particular actions can help us devise sound public policy. Let’s take a few examples. Hardin considers Bodo, who lived in the eleventh century and knew perhaps 80 people throughout his entire life. A moral theory that assesses actions made sense in Bodo’s world but not ours, says Hardin, who thinks that strategic and stochastic issues make action theory irrelevant. “In a simpler era,” he writes, “the focus on actions and act-kinds might have made approximate sense. Today it does not” (98).

Hardin also discusses scientist Edward Jenner, who deliberately inoculated a boy in 1796. Hardin writes, “This experiment presumably would not get past any ethics review board today” (87). The Kantian principle, he instructs us, is useless for understanding how to balance trade-offs among persons in stochastic contexts. Hardin does not consider whether Jenner could have recruited some adult to volunteer instead. The Kantian principle and common sense help us ask this question. We look for feasible alternatives and likely consequences, we look at relevant principles about autonomy and persons-as-ends; we assess individual actions, and we narrow down the options. These considerations all matter, and an ethics board convenes and determines an answer. This isn’t mechanics.

The oddity of Hardin’s project is that he throws all normative anchors to the wind, except one, basic rationality. He then focuses our attention solely on institutions and says that there is not much to say about how to organize society. In real life, however, when we witness the actions that are part of what institutions do, the anchors of moral thought come crashing back down into the sea. His theory offers no reflections why our moral world is, in fact, so determinate in practice. Hardin’s book simply has us spinning in the winds of individual interests, and nothing much comes of it.

We have a choice: we could go with his theoretical reflections, which take us nowhere (or at best almost nowhere), or go with the world that we actually observe and live in, which yields an enormous range of determinacy provided within our ethical and religious traditions of thought. These traditions sometimes offer conflicting and incomplete advice, but Hardin’s reflections leave us normatively, at the collective level, with nothing but mechanics. Just as choosing always-defect in prisoner’s dilemmas is something only a theorist could recommend (as Hardin tells us), so is arguing for the paucity of determinacies when we manage them every day.

This alternative to Hardin is moral theory as a practical pursuit. It helps us develop moral awareness. Evaluating acts and act-kinds are central for how we interpret the world and become aware of the relevant reflections. Take a recent example. We see a government deploying horrendous policies at home or abroad. We watch individual acts of repression, torture, beheadings, or worse. We watch Abu Ghraib. Is it true that the focus on actions and act-kinds make no sense for reforming and establishing sound public policy? This is why the pictures of Abu Ghraib are so powerful. We know: something went wrong. His two-stage theory does not admit this form of reasoning to reach determinacy, whether it be our repulsed responses to torture and murder and beheadings, or the more nuanced ethical complexities of medical research. This does not mean that we have resolved the issue, but we have anchors.

It is worth emphasizing two points of contrast between Hardin's approach of mechanical determinacy and this alternative. First, Hardin pursues theory through a reductionist approach (reducing all normative commitments to just one: individual interests), all in the name of common sense. The alternative is to collect and sort through a complex moral reality as people experience it, with the aim of developing reflections that assist and tailor that collection of experiences. These experiences fundamentally include our anchored moral assessments of actions. We have moral values, we give them meaning, and we let evidence bear on their proper interpretation. Hardin will have nothing of that empirical enterprise for reducing indeterminacy, at least in this book. In this sense, a focus on assessing actions, which Hardin explicitly rejects, invites a more empirically rich investigation than his abstract discussion of mutual advantage.

Second, significant focus on actions is more practical than Hardin's institutional perspective. We get up in the morning and pursue our lives within many overlapping institutions. How are we to understand our lives within them? Apparently, Hardin wants us to reason this way: imagine a choice between living in some institution or no institution. We see that selecting institutions is *ex ante* mutually advantageous. (I do not see this, but he thinks each and every one of us will or should.) We then accept utility comparisons for policy decisions on pragmatic grounds, even though we do not know what these trade-offs could mean, let alone what a more substantive justification could be (133). All of this abstruse theorizing is supposed to help me get up and get on with my day.

There may be compelling arguments to reduce away our complicated moral world into mechanics, but he hasn't given them. We live in a world with a rich moral vocabulary that captures our experiences, reduces our indeterminacies, and gets us to solutions, much of the time. Rather than let mechanics take over in the face of indeterminacy, we do better by studying these ethical traditions and studying how people give meaning to

their ideas of respect, liberty, justice, compassion, beneficence, generosity, and many others. These ideas, often centered on actions, are the bases for evaluating our political and economics institutions and for managing difficult trade-offs. They are the bases for understanding how to apply an idea of basic rationality and they provide answers to practical questions about social organization. If I had to place my bets, I would say that people in Bodo's time were keenly aware of strategic and stochastic indeterminacies, if not by those names, and that their world was not that much fundamentally different than our world today. Hardin's sole focus on mechanical determinacy can begin to make sense, perhaps, if you do not believe that.

Perhaps Hardin believes that the very notion of moral anchors and going beyond the normative commitments of basic rationality introduce its own indeterminacies. There is a point to that, and no single theory has all the answers, but one must live those practices and see what kinds of determinacies result. The project is to study how people live as much as how people theorize. This alternative is empirical and practical, and people seem to find considerable success at making decisions that guide social choice. Moral anchors about actions better explain how people deal with indeterminacies. They better explain how our social institutions come about. And they offer better reasons for making social choices.

In summary, Hardin's approach for understanding and responding to indeterminacy is by talking about institutions and mechanics. I argue that only by exploring the complexities of our moral experience can we begin to understand how people deal with indeterminacy.

Abraham again

Abraham's situation is a radical challenge. In *Fear and Trembling* Kierkegaard (1985) asks us to wrestle with actions contrary to reason, contrary to what we can intelligently communicate, to all anchors of ethical thought. Hardin's rational indeterminacy is about puzzles of social organization that emerge after we accept a clear anchor of normative thought, that a rational individual pursues more rather than less value. My argument is that our moral lives are filled with a multitude of anchors of ethical thought beyond basic rationality. They give us many determinate outcomes in our actual lives, despite any normative or explanatory theorizing by Hardin that tells us otherwise. To address the challenge of his book is to explore this richer normative world.

Even if my argument is correct, indeterminacy remains a deep puzzle. It is a puzzle for the story of Abraham, where we are asked to go beyond ethics. It is a puzzle for stochastic choice and strategic interaction, for we now go within ethics, and there is no simple determinate story to tell.

Even if the framework of Hardin's analysis is inadequate, his focus is compelling.

Steven Scalet

Binghamton University, SUNY

REFERENCES

- Frank, Robert. 2004. *What Price the Moral High Ground*. Princeton University Press
 Kierkegaard, Soren. 1985. *Fear and Trembling*. Penguin Books
 Kymlicka, Will. 2002. *Contemporary Political Philosophy: An Introduction*. Oxford University Press

doi:10.1017/S0266267105240671

The Difficulty of Tolerance, by T. M. Scanlon. Cambridge University Press, 2003, ix + 273 pages.

1. Scanlon says in the Introduction to *The Difficulty of Tolerance* that the 13 essays that comprise it "are concerned with the standards by which political, legal, and economic institutions should be assessed" (1). It thus represents an important complement and supplement to *What We Owe to Each Other* (1998), which is concerned mainly with the assessment of individual conduct. Subtle, insightful, and challenging, *The Difficulty of Tolerance* should be essential reading for anyone interested in the questions with which it engages.

Scanlon's concern with standards in the political domain proceeds at two main levels. The first level is *substantive*. This involves identifying what the relevant standards are and how they are to be weighed against one another. The second level is *justificatory*. This involves determining how these standards are themselves to be justified.¹

Scanlon is adamant throughout that well-being is one of the things that matters at both levels. However, this leaves open questions about *how* it matters and whether it is the *only* thing that matters. There are substantive questions about whether considerations of justice (rights, equality, fairness, the rule of law, and so on) are capable in certain circumstances of overriding well-being. And there are questions about whether the justification of some standards must make (non-trivial) reference to considerations other than

¹ There is also a third level, which involves applying these standards to the evaluation of particular policies and institutions. However, both for reasons of limited space and because Scanlon has much less to say about it, I shall largely ignore this third level in what follows.

well-being. Clearly, much rests upon what is meant by notions such as “well-being,” “justice,” and its constituents.

Because *The Difficulty of Tolerance* is a collection of essays, not a monograph, it proceeds in a somewhat piecemeal fashion, often in response to specific rather than general questions. Because it is a collection of essays written over a period of 30 years, there is no single stable position that endures throughout, but rather a series of evolving positions. Nonetheless, at the risk of oversimplifying things somewhat, I believe we can usefully distinguish three phases in Scanlon’s thinking about these questions, in which a more or less definite position can be discerned. In what follows, then, I shall attempt to briefly convey the main thrust of the three phases before raising in conclusion some objections to the position at which Scanlon arrives.

2. The first phase that I believe we can discern in Scanlon’s thinking (represented by essay 1) is what I shall call *the deontological phase*. At the level of justification, this involves the claim that “a legitimate government is one whose authority citizens can recognise while still regarding themselves as equal, autonomous, rational agents” (14f.). By autonomy Scanlon says that he means something “extremely weak” (15): “To regard himself as autonomous in the sense I have in mind a person must see himself as sovereign in deciding what to believe and in weighing competing reasons for action. He must apply to these tasks his own canons of rationality, and must recognise the need to defend his beliefs and decisions in accordance with these canons” (15).

Substantively, it involves the claim that there are substantive deontological constraints, at least with respect to violations of freedom of expression.² This is reflected in Scanlon’s endorsement of what he calls “the Millian Principle,” according to which:

There are certain harms which, although they would not occur but for certain acts of expression, nonetheless cannot be taken as part of a justification for legal restrictions of these acts. These harms are: (a) harms to certain individuals which consist of their coming to have false beliefs as a result of those acts of expression; (b) harmful consequences of acts performed as a result of those acts of expression, where the connection between the acts of expression and the subsequent harmful acts consists merely in the fact that the act of expression led the agents to believe (or increased their tendency to believe) these acts to be worth performing. (14)

The justification for the Millian Principle runs as follows. To be meaningful, the denial of (a) would require “conced[ing] to the state the

² Though it seems clear that Scanlon believes that a more extensive list of deontological constraints will be justified along the same lines.

right to decide that certain views were false and, once it had so decided, to prevent [any individual S] from hearing them advocated even if he might wish to" (17). Similarly, the denial of (b) would require "giv[ing] the state the right to deprive citizens of the grounds for arriving at an independent judgment as to whether the law should be obeyed" (18). Neither concession by S is compatible with S's "remaining autonomous" (17f.).

Scanlon's position in the deontological phase is susceptible to a number of objections, as Scanlon himself acknowledges. His justificatory account assigns way too much value to autonomy at the expense of other values. Moreover, weak as this conception of autonomy is, it is not so weak as not to require defense. Yet little or no defense for it is given. The Millian Principle suffers from the dual affliction of being false and relatively uninteresting. It is false since it implies that certain morally legitimate violations of freedom of expression – such as those involving "justified paternalism" (96) and/or those necessary to avoid truly monumental quantities of the two kinds of harms picked out by the Millian Principle (97) – are illegitimate. It is relatively uninteresting since it offers only a sufficient condition for one kind of morally illegitimate restriction on freedom of expression. Being incomplete, it is also extremely weak.³

3. The second phase that I believe we can discern (spanning essays 2–6) is what I shall call *the sophisticated rule consequentialist phase*. As its name suggests, this is an amalgam of a sophisticated consequentialist theory of justification and a theory of substance and principles that includes some non-consequentialist or deontological elements (such as rights). Scanlon acknowledges both aspects when he describes it as "a two-tier view: one that gives an important role to consequences in the justification and interpretation of rights but which takes rights seriously as placing limits on consequentialist reasoning at the level of casuistry" (26f.).

Like other consequentialists about justification, Scanlon holds that objects within the relevant domain of evaluation are ultimately "justified by appeal to the states of affairs that they promote" (35). Like other sophisticated consequentialists, he endorses a sophisticated theory of the good. This has two components. The first is a sophisticated conception of well-being as "ethically significant" and hence "objective." It is ethically significant in that ethical considerations cannot be divorced from the content of judgments about an individual's well-being. It is objective in that "a person's level of well-being . . . is independent of that person's tastes

³ Scanlon is not oblivious to this. As he admits: "On the basis of this principle alone, we could raise no objection against a government that banned all parades or demonstrations (they interfere with traffic), outlawed posters and handbills (too messy), banned public meeting of more than ten people (likely to be unruly), and restricted newspaper publication to one page per week (to save trees)" (21).

and interests" (72). Thus, "an appraisal could be correct even though it conflicted with the preferences of the individual in question, not only as he believes they are but even as they would be if rendered consistent, corrected for factual errors, etc." (72).⁴ The second component is the view that considerations of well-being – even sophisticated well-being – do not exhaust the considerations that determine the good. Scanlon also believes that there are other goods – for instance, fairness and equality – which are both "valuable in their own right" and "whose value does not rest on their being good things for particular individuals: fairness and equality do not represent ways in which individuals may be better off" (31–3).

At the level of substance, rather than embracing what we might call "institution consequentialism" or "policy consequentialism" according to which institutions or policies are the primary evaluanda,⁵ Scanlon embraces a version of "rule consequentialism" according to which rules or principles are the primary evaluanda. Institutions and policies are then evaluated in terms of whether or not they accord with the best available set of rules and principles. This allows him to incorporate elements of justice – for instance, rights – into his substantive first-order theory. Rights are "limitations and requirements . . . [that] must be imposed on policy decisions if we are to avoid results that would be unacceptable with respect to the considerations that are defined at the fundamental level" (99). This has the virtue of implying that, far from being "irrational" (6), rights are a crucial means of bringing about as much good as possible.

Nonetheless, there remain problems with Scanlon's position in the sophisticated rule consequentialist phase. One problem is that, arguably, sophisticated consequentialism is too sophisticated for its own good. It risks vindicating our intuitions about the right at the cost of violating our intuitions about the good; making mere axiology, upon which much of the appeal of consequentialism lies, impossible; and making consequentialism trivially true and therefore empty. A second set of problems concerns the structure of rule consequentialism. There is the fact that it seems guilty of "rule-worship." There is also the familiar dilemma concerning compliance.⁶ Though Scanlon has a response to the second – that involves pointing out that most of the rights with which we are concerned do

⁴ Scanlon is careful to point out, however, that objectivity is not the same as universality. It is perfectly conceivable that there may be variation across individuals and societies in respect of the objective well-being-makers. Nor does it foreclose the possibility that in many circumstances the promotion of an individual's well-being requires taking seriously the content of her subjective states, for instance satisfying her preferences. It is just to deny that the well-being *consists* in the satisfaction of those preferences.

⁵ I take institution consequentialism and policy consequentialism to be simply analogues of act consequentialism in the political context.

⁶ If valid rules are those that are best given actual compliance levels, then it will be hard to distinguish the view from act consequentialism. If, by contrast, idealizing assumptions are

not depend on a social practice of the right in question being generally observed (37) – he does not seem to have any plausible response to the first. A third problem concerns the apparent inability of his account to accommodate important intuitions that we have about the “separateness of persons.” If the Scanlon of the deontological phase took these intuitions too seriously, the Scanlon of the sophisticated rule consequentialist phase appears not to take them seriously enough.

4. The third and final phase (spanning essays 7–13) is what I shall call *the contractualist phase*. The most important and interesting shift occurs at the level of justification where Scanlon abandons sophisticated consequentialism in favour of his now well-known model of contractualism according to which standards, to be justified, must be such that no one could reasonably reject them “as a basis for informed, unforced general agreement” (132).

Scanlon’s contractualism has two main elements. The first is the notion of *reasonable rejectability*. To say that an individual N could reasonably reject a standard S is to say that N has sufficiently strong reasons to reject S. This is obviously meant to rule out cases of individuals rejecting standards or principles without adequate reason, for instance on the basis of sheer whim. Scanlon focuses on reasonable rejectability rather than reasonable acceptability on the grounds that certain intuitively invalid standards – for instance, standards that impose a severe hardship on some people – may, if those people “are particularly self-sacrificing,” nonetheless be standards that everyone could reasonably accept (133). Notice, moreover, that valid standards are those that *no one* could reasonably reject. Every individual holds, as it were, a power of reasonable moral veto. The existence of a single individual N such that N could reasonably reject S is sufficient to render S invalid.

The second element is what I shall call *the Scanlonian proviso*, that is, the idea that what matters is reasonable rejectability, given the motive of finding mutually justifiable principles. This is a crucial constraint on what sorts of rejections count as reasonable and unreasonable in the relevant sense. Consider for instance a standard that requires redistribution of the resources of the poor into the hands of the rich. It may be reasonable for Rupert Murdoch to accept this principle given the motive of advancing his material interests relative to others. However, the reasons that would make this true are of little or no *moral significance*. The Scanlonian proviso explains why this is so.

Well-being continues to play an important role in contractualist justification. Since we have reason to care about well-being, there will

made about compliance, then rights and other non-consequentialist first-order standards will increasingly come to seem “pointless” and irrelevant (37).

be many contexts in which the fact that a given standard will justify institutions or policies that would make us worse off provides a strong reason to reject the standard in question, given the aim of finding mutually acceptable standards (140). However, a negative contribution to one's net well-being does not provide the *only* reasonable grounds for rejecting a principle. Nor is well-being morally basic. Rather, its moral significance is determined by its being the sort of thing that we have reason to care about given the motive of finding mutually acceptable principles (184). Finally, it is not aggregate well-being but the well-being of particular individuals that figures as a salient input to contractualist justification.

Unsurprisingly, this contractualist theory is used to justify a kind of substantive pluralism. Well-being is one of the things that figures at the level of substance. But so too do justice and its various constituents. One important aspect of justice is rights. Rights are given substantive weight, not because they are required by autonomy, nor because doing so brings about good consequences, but because principles that assigned them inadequate weight could be reasonably rejected as a basis for informed, unforced general agreement. Tolerance provides another example. As Scanlon acknowledges, tolerance has major costs: costs for politics, and costs for individuals. To tolerate those with whom one disagrees is to risk, among other things, societal change of a sort that is undesirable from one's own point of view (195–200). On the other hand, there are also strong reasons to reject principles that would permit intolerant behavior. Foremost among these is the fact that intolerant behavior deprives us of the possibility of a certain meaningful relationship with others whereby we recognize "others as just as entitled as we are to contribute to the definition of our society" (193).⁷

5. This is not the place to attempt a full-scale evaluation of Scanlon's position in the contractualist phase. My own view (or prejudice perhaps) is that he has basically moved in the right direction. Nonetheless, I believe that many serious problems remain. In conclusion, I shall simply mention three of these.

The first is that it is unclear that the Scanlonian proviso has sufficiently determinate content to rule out many principles. According to Scanlon, in order for a principle P to be ruled out, it must be the case that there is some individual N such that N has sufficiently good reason to reject P, given the aim of finding principles that others, similarly motivated, could not reasonably reject. But how might we begin to determine whether or not this condition is satisfied? It seems to me that we do not have enough information to do so. Consider a principle that requires that

⁷ Other aspects of justice that, according to Scanlon, are supported by contractualism include the rule of law (essay 12) and consent (essay 13).

criminal laws treat defendants as innocent until proven guilty. Now take an affected citizen Gertrude. Does Gertrude have sufficient reason to reject this principle, given the aim of finding principles that others, similarly motivated, could not reasonably reject? It is not obvious to me that there is a determinate answer to the question. This is not altogether surprising, given that the charge of indeterminacy has often been filed against Kant's universalisability test of which the Scanlonian proviso is one version. Moreover, it would also offer one explanation for why many critics of Scanlon have thought that his contractualist formula is explanatorily otiose and redundant, or at least incomplete. For, if this objection is correct, additional moral content would have to be packed into the contractual situation to render the formula determinate.

A second problem is what I shall call the problem of scope. The problem of scope concerns the scope of beings whose interests are given moral weight and the manner in which this is done. Actual, mentally normal, adult humans are not the only beings who stand to be affected by policies and institutions. So too can non-human animals, the severely mentally handicapped, children, fetuses, past and future generations, and so on. Scanlon's contractualism may imply that the interests of these beings carry some substantive weight. However, even if it does, it does not seem to do so in the right way. For the justification of standards does not make essential and direct reference to the beings themselves.

The problem of scope is not, of course, an original problem. But it is a source of genuine amazement to me that Scanlon has paid it so little attention. Scanlon's "solution" in *What We Owe to Each Other* is to insist that contractualism is only meant to be about one part of morality. But this is altogether inadequate. A test of a good theory is surely how it deals with particularly difficult phenomena. It is highly dubious to suggest that one's theory is not supposed to cover those phenomena with respect to which it obviously falls short. Imagine that a consequentialist were to say that her model of consequentialism is not supposed to cover especially demanding duties, such as those to persons in the third world.

The third problem concerns the applicability of Scanlon's model of contractualism to the political context. Recall that this model is presented in "Contractualism and Utilitarianism" primarily as a theory of individual conduct, not as a theory of policies and institutions. How then is Scanlon's model of contractualism to be interpreted as a theory of the latter? As far as I can see, there are two possibilities. The first is that it is meant to give expression to the same mode of assessment that is relevant in the case of individual conduct (the moral mode) simply applied to a different object of assessment (policies and institutions). The problem is that, as it stands, his contractualism assigns to all persons equal power to determine what standards of assessment are valid and invalid. Yet, intuitively, this power should rest wholly or at any rate mainly in the hands of the citizens

and residents of the polity in which those policies and institutions are being implemented or maintained. This leads me to think that Scanlon must have in mind a second possibility, which is that his contractualism is meant to be giving expression to a distinct though related mode of justification – what we might call political justification. If this is right, to make sense, his contractualist formula must be amended to make clear that the validity of political standards is a matter of their being such that no *citizen or resident* could reasonably reject them given the aim of finding standards that other *citizens or residents*, similarly motivated, could not reasonably reject. This raises interesting complications concerning the relation between the moral and political modes of assessment, but I cannot address those complications here.

Nicholas Southwood

Australian National University

doi:10.1017/S0266267105250678

Hayek's Challenge – An Intellectual Biography of F. A. Hayek, by Bruce Caldwell. University of Chicago Press, 2004, xi + 489 pages.

Bruce Caldwell is a widely acknowledged expert in the history of economic thought as well as in economic methodology and, as he notes in the introduction to *Hayek's Challenge*, much of his research interest over the past two decades has been focused on the work of F.A. Hayek. He has edited two volumes of *The Collected Works of F.A. Hayek* (vol. 9, *Contra Keynes and Cambridge*, 1995; vol. 10, *Socialism and War*, 1997), and he now serves as the general editor of the series. The combination of interest and expertise he brings to the task has enabled him to write an impressive intellectual biography of Hayek that, with its specific focus on methodological issues, is not only highly instructive about its particular subject but also most relevant for the broader issue of how economics as a scientific enterprise ought to position itself.

The title, *Hayek's Challenge*, is meant as a reference to the challenges that Hayek faced as someone whose economic and political positions were often “completely out of sync with those of the rest of the intelligentsia” (3), as well as a reference to the challenges that Hayek presents “to those who try to interpret his thought” (4). Yet, it also indicates a major underlying motivation of Caldwell's book, namely to challenge dominant trends in modern mainstream economics, trends that are in fundamental contradiction to Hayek's vision of what the social sciences can reasonably

be expected to accomplish and of how social phenomena may best be studied.

As the subtitle indicates, *Hayek's Challenge* is an *intellectual* biography. It does not, and is not meant to, tell us much about Hayek's personal life (readers interested in biographical information may turn to Alan Ebenstein's *Friedrich Hayek: A Biography*, University of Chicago Press, 2001). Even as an intellectual biography it is, deliberately, selective, focusing mainly on those parts of Hayek's work that are of particular relevance for Caldwell's principal concern, the development of Hayek's methodological views and, specifically, the lessons that Hayek's thought holds for those modern day economists who believe that mathematical modeling and econometric studies constitute the hallmark of a "science" of economics. It is, in fact, its selective focus that makes *Hayek's Challenge* a particularly valuable contribution not only to the history of economic thought but, equally, to current debates on the state of the discipline and on the directions in which it ought to develop.

Apart from a brief introduction and four appendices on particular side issues (407–38) Caldwell's book is divided into three parts. Part I, "The Austrian School and its Opponents – Historicists, Socialists, and Positivists" (17–130), discusses the intellectual background from which, and the intellectual environment in which, Hayek's thought emerged. In five chapters, Caldwell traces the evolution of the Austrian School, from Carl Menger's foundational contribution, to his and his followers' clash with the German Historical School, and to Ludwig von Mises' critique of socialist and positivist currents of thought. On the way, Caldwell provides a wealth of contextual information, e.g., on Gustav Schmoller and the "Methodenstreit," on the emergence, the influence, and eventual decline of the German Historical School, on Max Weber, on the peculiar place of Josef Schumpeter in the Austrian tradition and on Otto Neurath's socialism and positivism. As a prelude to the book's proper subject, the study of Hayek's own thought, this part requires some patience on the reader's part. Yet such patience pays off well, especially for readers less familiar with the Austrian School's approach to economics. It represents not only a most instructive account of the emergence of the intellectual tradition that Hayek carried forward. More importantly, it provides an excellent introduction to the issues around which Hayek's work centers.

In the second and main part of his book, titled *Hayek's Journey* (133–319),¹ Caldwell traces the development of Hayek's thought, his "intellectual voyage." The first three of the seven chapters in this part cover the beginnings of Hayek's academic career in Vienna, paying particular attention to methodological aspects of his early writings on

¹ Incidentally, there is a noteworthy coincidence with the title Alan Ebenstein has chosen for his second book on Hayek, *Hayek's Journey – The Mind of Friedrich Hayek* (2003).

monetary issues and on business cycle theory, and his move to the London School of Economics, focusing on his relation to Lionel Robbins, his battle with Keynes, and on methodological debates, largely initiated by Robbins' *Essay on the Nature and Significance of Economic Science* (1935), that figured prominently in the intellectual environment to which Hayek was exposed in the 1930s. The other four chapters, titled "Economics and Knowledge' and Hayek's Transformation," "The Abuse of Reason Project," "Individualism and the Sensory Order," and "Rules, Orders, and Evolution," are the key chapters in Caldwell's account of "Hayek's Journey."

In "Economics and Knowledge' and Hayek's transformation" Caldwell draws particular attention to the 1937 article "Economics and Knowledge," the published version of the Presidential address that Hayek gave to the London Economic Club in November 1936 and that, as he later noted in retrospect, marked the beginning of an intellectual journey that led him from "technical economics into all kinds of questions usually regarded as philosophical" (205). In "Economics and Knowledge" Hayek clearly stated the issue, commonly referred to as the "knowledge problem," that was to become the central concern of his research efforts in the following decades. If we take seriously the subjective nature and incurable limitations of what humans know about the world, then, and this was Hayek's argument, we must conclude that standard economic analysis, by assuming "that the same objective data, the same knowledge, are *given* to all agents" (207–8), simply bypasses the very questions that need to be answered. Namely, how knowledge about the world is acquired by the human mind, how the dispersed and fragmented pieces of knowledge are communicated among individuals, and how individuals who are guided in their actions by subjective knowledge are able to successfully coordinate their activities. As Caldwell shows, the knowledge argument, first stated in "Economics and Knowledge" and reemphasized later in "The Use of Knowledge in Society" (1945) and "Competition as a Discovery Procedure" (1968), was to become the centerpiece of Hayek's system of thought. It is as central to his position in the socialist calculation debate as it is to his rejection of von Mises' apriorism and his critique of the equilibrium construct of mainstream economic theory.

The chapter on "The Abuse of Reason Project" takes a closer look at the first major research endeavor that grew out of Hayek's concern with the knowledge problem, his critical analysis of "the abuse and decline of reason in modern time" (240), an "abuse and decline" that he diagnosed as resulting from a "pretence of knowledge," an exaggerated, "scientific" belief in what science can achieve in social matters. The chapter begins with an instructive account of the anti-liberal, pro-planning mindset that dominated the intellectual landscape of 1930s Britain, a mindset against which Hayek took a pronounced stand, first with a

pamphlet on "Freedom and the Economic System" (1939) and then, in expanded form, with "The Road to Serfdom" (1944). As Caldwell argues, in Hayek's view "the widespread enthusiasm for socialism and planning among virtually all the intelligentsia" could not be explained as a mere response to currently perceived economic problems but "had its origins in ideas found in the past" (253–4). This perception motivated his inquiry into the historical roots of the modern scientific mentality, resulting in a sequel of articles: "The Counter-Revolution of Science" (1941), "Scientism and the Study of Society" (1942–44), "Individualism: True and False" (1946), and "Comte and Hegel" (1951). Caldwell attributes particular significance to the "Scientism" essay as a "transition piece" (341) in Hayek's intellectual journey that contains the germs of two separate, but gradually intertwining, research projects that Hayek worked on between 1945 and 1960. On the one hand, an inquiry into the working of the human mind, into how individuals' subjective knowledge is formed, a project that resulted, in particular, in "The Sensory Order: An Inquiry into the Foundations of Theoretical Psychology" (1952). And, on the other hand, an inquiry into how social institutions condition the communication, coordination, and discovery of knowledge, a project that resulted, in particular, in Hayek's portrayal and defense of a classical liberal order in "The Constitution of Liberty" (1960). In "Individualism and the Sensory Order," Caldwell discusses the significance of Hayek's excursion into theoretical psychology in the context of his overall project, i.e., the endeavor to advance our theoretical understanding of how the social world works. Caldwell points out that, even though "Hayek never directly linked his economics to his theory of the mind" (277), there are important systematic links between the two, a fact that Hayek commented upon in retrospect when he said about *The Sensory Order*: "The work on it has helped me greatly to clear my mind on much that is relevant to social theory. My conception of evolution, of a spontaneous order and of the methods and limits of our endeavors to explain complex phenomena have been formed largely in the course of the work on that book" (278–9). Apart from the role that *The Sensory Order* plays in specifying the behavioral foundation of Hayek's social theory, Caldwell specifically draws attention to the analogies between Hayek's interpretation of the mind as "a vast network of interconnected neurons" (265) – an interpretation that, as Caldwell notes, "has many affinities with recent connectionist, or neural network, theories of the mind" (343) – and his theory of the formation of the market order. A common feature of both is the notion of a spontaneously formed complex order that we can seek to understand in terms of its general working principles, but that, because of the complexity of its concrete manifestations, we will never be able to fully explain in its details.

In "Rules, Orders, and Evolution," the chapter that concludes his account of Hayek's intellectual journey, Caldwell focuses on Hayek's ideas,

elaborated in *The Constitution of Liberty* (1960), in *Law, Legislation and Liberty* (1973, 1976, 1979) as well as in a number of essays, “about rule-following behavior; about the complex orders that sometimes spontaneously result when rules are followed; and about the ways that rules and orders develop and change, or evolve, through time” (288). Caldwell’s particular emphasis here is on the change in the ways in which Hayek sought to characterize the subject matter of economics and other social sciences, and to separate what he regarded as an adequate scientific approach to the study of society from scientific pretensions. As Caldwell observes, while Hayek originally, in the 1940s with his “Scientism” essay, emphasized the dividing line between the natural and the social sciences, during the 1950s he came to “the conclusion that the basic dividing line among all the sciences was between those that studied simple and those that studied complex phenomena” (304). By thus drawing a distinction that cuts across the traditional division between the natural and the social sciences, placing the social sciences in the same camp as, for instance, evolutionary biology, Hayek made it unambiguously clear that his critique of “scientism” in its various incarnations was not meant at all as an argument against a naturalistic, scientific approach to the study of society. As Caldwell puts it: “With this new distinction, it is clear that the social sciences are fully *scientific*; it is just that they are among those sciences that study complex phenomena” (305). With other sciences of complex phenomena, including those traditionally classified as “natural sciences,” they share the fate that, due to the complexity of their subject matter, their ambitions must typically be confined to “explanations of the principle” and to “pattern predictions.” Such a limitation, Hayek insists, is not due “to the immaturity of the sciences concerned . . . a provisional and transitory state . . . which they are bound to overcome sooner or later” (370), nor does it make sciences of complex phenomena any less “scientific” than the sciences that, like classical mechanics, study relatively simple phenomena. In fact, according to Hayek, as they advance we should expect more and more fields of inquiry to share in this limitation: “As the advance of the sciences penetrates further and further into more complex phenomena, theories which merely provide explanations of the principle . . . may become more and more the rule than the exception” (303).

“Hayek’s Challenge” is the title Caldwell has chosen not only for his book but also for its third and concluding part (323–405). In two chapters, “Journey’s End – Hayek’s Multiple Challenges” and “Epilogue: A Meditation on Twentieth-Century Economics,” he offers a summary assessment of Hayek’s methodological legacy and draws some conclusions that his account of Hayek’s intellectual journey implies for how one judges the direction that modern economics has taken. Caldwell’s principal conclusion is that, if Hayek’s view of the nature of the social sciences in general, and economics in particular, was right, “then some of the

directions taken by the discipline in the twentieth century have been wrong" (13). Distinguishing between what he calls "basic economic reasoning" and "the 'analytical constructions' of... formal economic theory" (330–1), Caldwell notes that the issue is not about basic economic reasoning, i.e., the kind of arguments economists commonly employ when they talk about the predictable effects of such things as minimum wage laws or rent control. "Basic economic reasoning," Caldwell argues, "is what is fundamental in economics" (336), and if it appears to work reasonably well this is because "it does not depend on the unrealistic assumptions that, for many decades, economists made in their formal models" (336). The issue is about the insights into their subject matter that economists have been able to provide with their "ever more sophisticated models" (*ibid.*). It is about the huge discrepancy between, on the one side, the pretentious formal modeling apparatus and the hyper-technical econometric methods that are employed and, on the other side, the poverty of the substantive insights into economic reality that are actually generated.

In Caldwell's assessment it is not only the poor performance record of formal mathematical modeling in economics (if one measures performance not in terms of the sophistication of its instruments but in terms of its substantive insights) that speaks in favor of Hayek's alternative methodological vision. Hayek's distinctive vision of economics as a science of complex phenomena has found strong support in recent work in such fields as complexity theory and the study of self-organizing complex adaptive systems, work that has striking parallels with Hayekian ideas. As Caldwell comments: "Perhaps the most heartening aspect of these new developments is that others in the scientific community are now coming to recognize, as Hayek did, that the nature of the phenomena that economists and other social scientists study may require approaches quite different from those utilized during earlier eras" (368).

Caldwell's *Hayek's Challenge* is a book with a message. It has not been written as a detached contribution to the history of economic thought. Quite apparently, the account that it provides of Hayek's intellectual journey is motivated by the conviction that the theoretical outlook that Hayek came to develop over his lifetime holds important lessons for an economics profession that, in Caldwell's judgment, has been misguided by a false understanding of how it is to prove its "scientific" status. In tracing the development of Hayek's thought as he struggled with "scientistic" misconceptions of what social science can achieve, Caldwell alerts the economics profession to the fact that its obsession with mathematical virtuosity may reflect a narrow-minded, antiquated understanding of science as well as a failure to recognize the limits that economics as a science of complex phenomena faces. In insisting on these limits Hayek did not mean to suggest at all, as Caldwell shows, that economics give up its ambition to be a *scientific* enterprise. To the contrary, it was his

ambition “to defeat his scientific opponents with what he thought were truly scientific arguments” (260).

In addition to its value as an important contribution to the history of economic thought, Caldwell’s book may be of even greater significance as a contribution to the ongoing debate on the current state and the future direction of the discipline of economics. One should wish that *Hayek’s Challenge* finds attention not only among those who share Caldwell’s general outlook, but also among those at whom the challenge is directed. With his account and methodological assessment of Hayek’s intellectual journey Caldwell has not only added a forceful battery of arguments to the growing choir of critics that voice discontent with the state of the profession, he has also shown that Hayek’s thought suggests a promising alternative research program for economic science, a program that is much more in line with recent developments in the study of complex adaptive systems than the scientific formalism of mainstream economics.

Viktor Vanberg

University of Freiburg