FOREWORD TO THE THIRD EDITION

Eric S. Maskin

Kenneth Arrow is a towering figure in economics and the social sciences more generally. And so I felt extraordinarily honored when Yale University Press asked me to write a foreword to a new edition of his landmark monograph *Social Choice and Individual Values.* Ken also happens to be my former teacher and a mentor and role model—facts that made the invitation a great personal pleasure as well, and an excuse for some nostalgia.

I first met Ken when I was an undergraduate at Harvard in the early 1970s. I was a math major but somehow wandered into his graduate course on information economics. It was a hodgepodge of some of the cutting-edge topics he was thinking about then—e.g., mechanism design, adverse selection, and communication in organizations—and proved to be utterly fascinating. I can't say that Ken—despite his interest in the theory of organizations—was the most organized of teachers. His lectures sounded improvised, largely because they were: he apparently decided what to talk about on his way over to the classroom (and sometimes not even then). On one unusual occasion, he prepared a lecture in advance—on a highly technical result called the Gibbard-Satterthwaite theorem (not yet published at the time)—but then forgot to bring his notes. He worked out a new and detailed proof of the theorem for us on the spot.

One feature that made the course so gripping was Ken's ability to pack so much into so little time. This was partly because he talks unusually fast. But even Ken's rapid-fire speaking style couldn't keep pace with his lightning-quick mind. So his presentations had an elliptical quality—he would leave out the ends of sentences so that he could race ahead to the next thought. His listeners would have to stay alert to fill in the missing words.

Another thing that struck us students was the breadth of Ken's knowledge, which extended far beyond economics. He would be lecturing on some technical point when an apt quotation from Maimonides or an analogy from thermodynamics would occur to him. Ken evidently knew more on virtually any subject than any of

the rest of us. There was a story at Harvard that a group of junior faculty once concocted a plan by which they could finally appear to outshine their erudite senior colleague. They read up on the most arcane topic they could think of: the breeding habits of gray whales. On the appointed day they gathered in the coffee room and waited for Ken to come in. Then they started talking about the elaborate theory of a marine biologist named Turner on how gray whales find their way back to the same breeding spot year after year. Ken was silent . . . they had him at last! With a delicious sense of triumph, they continued to discuss Turner, while Ken looked increasingly perplexed. Finally, he couldn't hold back: "But I thought Turner's theory was discredited by Spenser, who showed that the supposed homing mechanism couldn't possibly work."

With its publication in 1951, *Social Choice and Individual Values* initiated the modern theory of social choice, the study of how a society should choose among its various options based on the preferences of the individual members of society. There had been sporadic literature on the subject before Arrow, going back (at least) to Jean-Charles Borda and the Marquis de Condorcet in the late eighteenth century. But the earlier essays lacked the generality and power of Arrow's approach, and the subject did not really take off until *Social Choice.* But take off it did: by the time the second edition was published, in 1963, there were already several hundred works building on the book. A recent count on Google Scholar turned up over ten thousand citations.

Let me try to explain why the monograph has turned out to be so immensely influential—of interest to political scientists, sociologists, lawyers, and philosophers as well as economists. First, Arrow's abstract formulation of the social choice problem makes it very widely applicable. He begins with a society and a set of social alternatives (the different possible options from which society must choose), which, depending on the context, could be almost anything. For example, in a setting where a town is considering whether or not to build a bridge across the local river, "society" comprises the citizens of the town, and the social alternative set consists of just two options: "build the bridge" or "don't build it." In a context of pure distribution where there is a jug of milk and a plate of cookies to be divided among a group of children, the children are the society, and the social alternative set consists of the different ways the milk and cookies could be allocated to them. In a setting where a committee is interested in electing a chairman, society is the committee, and the social alternatives are the various candidates for the chairmanship.

Arrow's definition of a *social welfare function* (SWF) is also very general. An SWF is any rule for determining *society's* preferences

over the social alternative set on the basis of the preferences of the *individual members*. More precisely, because individuals' preferences might not be known in advance, the SWF is a function: it must determine social preferences for every different configuration of preferences that individuals could have, i.e., for every *profile* of possible preferences.

The most famous example of an SWF is probably *majority rule*, which Condorcet himself particularly advocated for elections. Under majority rule, alternative a is socially preferred to b for a given profile if a majority of individuals prefer a to b for that profile.

A second reason for the great impact of Arrow's monograph is the powerful and unexpected Impossibility Theorem that constitutes its central finding. Arrow discovered that there is *no* SWF that satisfies all of a few natural and seemingly undemanding conditions. These conditions are:

Unrestricted Domain (U): the SWF must determine social preferences for *all logically possible* profiles; in other words, there are no limitations on the preferences that individuals might have.

Pareto Property (P): if all individuals prefer alternative a to b, then a must be socially preferred to b.

Independence of Irrelevant Alternatives (IIA): if there are two profiles and each individual ranks alternatives a and b the same way in both of them, then the social preference between a and b must also be the same for both. In other words, the social preferences between a and b depend only on individuals' preferences between *these two* alternatives—and not on preferences involving some third alternative.

Nondictatorship (ND): there does not exist a member of society who always gets his way, in the sense that, for any profile, the social preferences coincide with his preferences.

Transitivity (T): for any profile, if a is socially preferred to b and b is socially preferred to c, then a must be socially preferred to c.

The Impossibility Theorem establishes that if there are at least three alternatives in the set of social alternatives, then there is no SWF that satisfies all of U, P, IIA, ND, and T.

It is worth noting why majority rule is not a counterexample to the theorem. As Condorcet himself pointed out, majority rule violates T. Suppose there are three alternatives a, b, and c, and consider a profile in which 35 percent of individuals prefer a to b and b to c, 33 percent

prefer b to c and c to a, and the remaining 32 percent prefer c to a and a to b. Then, 67 percent of individuals prefer a to b, and so a is socially preferred to b. Similarly, b is socially preferred to c (because 68 percent of individuals prefer b). But c is socially preferred to a (65 percent prefer c), and so T is violated.

Much of the literature stemming from *Social Choice* amounts to trying to avoid the Impossibility Theorem by relaxing one or more of the Arrow conditions. Condition ND is already so mild that relaxing it further seems pointless. The same might be said for condition P (but there is interesting work in this direction by Robert Wilson). By contrast, condition T is judged by some (see, in particular, James Buchanan) to be too strong a requirement. Nevertheless, relaxations of T turn out not to take us very far away from impossibility.

But relaxing condition U—which amounts to restricting the domain of preferences that an individual might have—has proved very fruitful. Even before Arrow, Duncan Black showed that in some applications, a natural restriction on preferences ensures that majority rule will satisfy T (in other words, it rules out the sort of T-violating example I gave above). In subsequent work, Amartya Sen characterized all possible restrictions under which majority rule is transitive, while Ehud Kalai and Eitan Muller and, independently, I characterized all possible restrictions under which there exists *some* SWF satisfying Arrow's remaining conditions. In recent work, Partha Dasgupta and I have shown that there is a sense in which majority rule satisfies Arrow's conditions (somewhat strengthened) for more domains of preferences than any other SWF.

Weakenings of IIA—the most controversial of Arrow's conditions—have spawned the biggest literature of all. As formulated, IIA rules out interpersonal comparisons: if in a two-member society, individual 1 prefers a to b and individual 2 prefers b to a in each of two profiles, then IIA requires that the social preferences between a and b be the *same* for both profiles—despite the possibility that in one profile, individual 1 has a strong preference for a over b and 2 only slightly prefers b to a, and the opposite is true for the other profile. Hence, in the extensive literature on interpersonal comparisons, IIA is weakened so that differences in preference intensity across individuals can be reflected in the SWF (although there are serious difficulties with trying to take account of such intensity differences when conducting elections). Recently, there has also been work developing other important relaxations of IIA, notably that of Marc Fleurbaey and François Maniquet, and also of Michel Balinski and Rida Laraki.

IIA also lies behind my final explanation for *Social Choice*'s great sway: the condition is intimately connected with the vast literature on mechanism design theory. Mechanism design was developed

to overcome an important obstacle to social choice, namely, that individuals' preferences—the inputs to the SWF—may not be publicly known. An SWF is said to be *implementable* if it is possible to design a *mechanism*—a procedure or game for individuals to follow—that leads to the choice of the top-ranked alternative according to the SWF (despite this lack of public knowledge). Arrow's monograph was a crucial foundation for mechanism design, because it turned out that conditions closely related to IIA are the key to an SWF's being implementable.

A book's importance can be crudely gauged by how many other works cite it. But perhaps a better measure is its longevity: how long it continues to inspire new work. By that criterion, *Social Choice and Individual Values* is an amazing success: having passed its sixtieth birthday, it continues to generate a steady stream of original research. I suspect that the same will be true when it reaches one hundred.