Democracy Defended

Gerry Mackie

"Arrow General Possibility Theorem"

2003
CAMBRIDGE
UNIVERSITY PRESS

4 The Arrow general possibility theorem

Democratic voting as meaningless

First, irrationalism claims that voting is arbitrary. Second, irrationalism claims that voting is meaningless: even if a voting method survives the first claim as fair, it is yet meaningless, because: (a) the outcome of voting is manipulable; *and* (b) we cannot know that manipulation occurred since again there is not enough information available from the data of voting to know the preferences underlying choices expressed in voting. The second claim of meaninglessness presents and interprets results of social choice theory. We have already treated (in Chapter 2) the crucial premise that preferences cannot be known from choices. Now, we will begin examination of the premise that voting is manipulable. The premise of manipulability is derived from the possibility of majority cycling as shown by Arrow (1963/1951), the possibility of strategic voting as shown by Gibbard (1973) and Satterthwaite (1975), the possibility of agenda control as shown by McKelvey (1976) and Schofield (1978), and finally the strategic introduction of new issues and dimensions (Riker 1982). Over the next three chapters, we discuss the Arrow theorem. In this chapter, we review the origins of the Arrow theorem in the ordinalist revolution in economics, and distinguish social choice as welfare economics from social choice as voting theory. Next, we present the contents of the Arrow theorem, followed by discussion of claims of its empirical relevance by Arrow and Riker. Then we review all studies found on the question of the frequency of cycles, and conclude that the incidence of cycles is rare. Finally, we begin review of justifications of the conditions of the theorem.

The origins of social choice theory

Classical economics emphasized the remarkable coordinating power of markets, and suggested that a policy of laissez-faire would best advance the wealth of nations. Bentham and the utilitarians embarked on

a wholesale program of practical reform, which overturned traditional policies, and enacted new policies, many laissez-faire, intended to do the greatest good for the greatest number. Economists were still bedeviled by the value paradox – why should useful water cost less than useless diamonds? The marginalist revolution of 1871–1874 solved that problem. The price of something is related to its marginal utility, not its total utility – water has great total utility, but in normal circumstances one more unit has little marginal utility – and consumer satisfaction is maximized when the ratios of marginal utility to price are equal for each good. Further, goods have a declining marginal utility: after some point, each additional increment is worth less than the prior increment. The intersection of utilitarianism with marginal analysis in Marshall's neoclassical economics yielded a conclusion distasteful to those fond of laissez-faire: if for each person there is a declining marginal utility of money, then it would increase overall social welfare if money were taken from the rich and given to the poor, up to the point equalizing marginal utility of each person in society. Efficiency would best be achieved by equality.

Utilitarianism had assumed cardinal and interpersonally comparable utilities, and the utilitarian philosopher proposed pursuit of the greatest good for the greatest number, that society should maximize the total sum of utility. Cardinal utility counts it as meaningful to say that I want a holiday in Andalusia five times more than I do a holiday in Buffalo; interpersonal comparability counts it as meaningful to say that Paul likes playing the guitar more than Matthew likes doing the dishes. Although interpersonal comparisons of welfare are common in daily life, and in my view are quite meaningful, they are always open to skeptical attack, and it was once thought that there are insurmountable difficulties in devising satisfactory formal representations of such comparisons. Meanwhile, economists found that they could restate the basic propositions of market economics in terms of ordinal and noncomparable utility. The advantage is that these are less demanding assumptions, and there are not so many formal and conceptual problems as there are with cardinal utility, comparable utility, or both. Ordinal utility considers only the order of ranking of alternatives; in the ordinal framework I cannot say that I like Andalusia five times better than I do Buffalo, only that I like Andalusia better than Buffalo. Further, a notion such as that society should maximize the total sum of utility is not possible within an ordinal and noncomparable framework; for one thing, you can't add what you can't compare.

The chief ideologist of the ordinalist revolution was Lionel Robbins (1937/1932), who wanted to establish economics as a "Science," and to

distinguish it from "Ethics." The purpose of a distinction between facts and values, is, according to Iris Murdoch (1992, 25):

> to *segregate* value in order to keep it pure and untainted, nor derived from or mixed with empirical facts. This move however, in time and as interpreted, may result in a diminished, even perfunctory account of morality, leading (with the increasing prestige of science) to a marginilisation of 'the ethical.' This originally well-intentioned segregation then ignores an obvious and important fact of human existence, the way in which almost all our concepts and activities involve evaluation.

I agree that it is quite important to distinguish fact from value, but notice that the claim I just made involves an assertion of fact and an expression of value. Too often, the discourse which states that value claims are *nonscientific* (in the descriptive sense) is twisted into an insinuation that value claims are *unscientific* (in the evaluative sense), or merely arbitrary expressions ungrounded in reason. Robbins tends to do this himself, in imagining that a committee made up of an economist, Bentham, Buddha, Lenin, and the head of US Steel would be unable to agree on the ethics of usury, but that the same committee would be able to agree on the facts of the economic consequences of anti-usury legislation (1937/1932, 150–151). Values are arbitrary, according to Robbins (150): "If we disagree about ends it is a case of thy blood or mine – or live and let live according to the importance of the difference or the strength of the opponents." At the same time he does not seem to be aware that his recommendations as to what should count as science are matters of evaluation, not of fact.

Robbins's second move was to allege that the claims of the reigning material-welfare school in economics were not scientific but ethical, and not just ethical, but arbitrary because ethical. The material welfare school held, according to Robbins, that it is possible to compare the utility or satisfaction of one person to another person. But such comparisons are not needed in modern economic theory, he wrote; the comparison is essentially normative and has no place in pure science.

> *There is no means of testing the magnitude of A's satisfaction as compared with B's.* If we tested the state of their blood-streams, that would be a test of blood, not satisfaction . . . There is no way of comparing the satisfactions of different people . . . In Western democracies we assume for certain purposes that men in similar circumstances are capable of equal satisfactions . . . although it may be convenient to assume this, there is no way of proving that the assumption rests on ascertainable fact. And, indeed, if the representative of some other civilization were to assure us that we were wrong, that members of his caste (or race) were capable of experiencing ten times as much satisfaction from given incomes as members of an inferior caste (or an "inferior" race), we could not refute him . . . we could not show that he was wrong in any objective sense, any more than we could show that we are right. (Robbins 1937/1932, 139–140)

From this he concludes that the recommendation by the material-welfare school for equalization of incomes is unscientific. Notice that Robbins's doctrine is radical skepticism rather than mere behaviorism: *any* objective correlate of satisfaction is prohibited. If we could measure some chemical in the blood (or these days, study an image of the brain's activation), that would not do; we would be measuring blood, not satisfaction. Robbins's objection seems to me to be one of postured philosophy rather than of ordinary science – science frequently estimates unobserved variables by way of indirect measures, without calling into doubt the theoretical usefulness of the unobserved entity. I agree with Robbins, though, that in any case the next step of saying that satisfaction or some other measure *should* be equalized – or *any* other policy recommendation, including that it should not be equalized – is a normative question.

Cooter and Rappoport (1984) argue that Robbins and his followers misconstrued the material-welfare school they superseded. Generally, the material-welfare school understood that ordinalism is sufficient for *market* economics, and that it may not be possible to compare the *satisfactions* of any *two* individuals. Where they differed was in judging that ordinalism is not sufficient for *welfare* economics, and in thinking it is possible to compare the *needs* of *representative* persons: "If people typically desire what they need, and if needs are more urgent when people are poor, then it follows that additional income is more useful to the poor than the rich" (Cooter and Rappoport 1984, 517). In their social-welfare calculation, goods were evaluated objectively, by whether they contributed to a person's physical well-being, they distinguished necessities from comforts from luxuries, and they measured variation among individuals in the supply of health, food, housing, clothing, and money. The welfare economist need not be confined to the equalization of satisfaction, a mental state; the welfare economist could have an objective theory of the human good, justified in its own right, and not solely because it correlates with the desire-satisfaction of the typical individual. Sen (1999; see also 1982) says that the rejection of interpersonal comparisons in welfare economics was based on interpreting them entirely as comparisons of mental states. He argues that, "even with such mental state comparisons, the case for unqualified rejection is difficult to sustain" (1999, 358). He continues that such comparisons need not be based only on mental states, but might directly be based on incomes, or commodity bundles, or resources more generally, and Sen's theoretical and applied work in this area demands attention.[1] To conclude, is giving food to the hungry better than giving

opera tickets to the bored (Cooter and Rappoport, 1984, 519)? I agree with Robbins that such is an ethical claim, but disagree that it is arbitrary because ethical.

As economists completed their ordinalist revolution in the 1930s, they sought in welfare economics to devise an ordinalist replacement for the utilitarian formula. They had the Pareto criterion, that x is better than y if every individual ranks x higher than y, but that criterion is radically incomplete: a policy change that helps a million people but hurts one is not a Pareto improvement, and further there is no way to choose among a multitude of Pareto-superior states. As it happens, voluntary market exchange satisfies the Pareto criterion, but collective choice short of unanimity does not, and thus any political distribution of endowments other than the inherited status quo is off limits. Notice that most voting schemes are ordinalist – voters are asked to rank-order alternatives, not to state intensities on some scale. There was a minor, almost forgotten, current of analytic consideration of alternative voting rules, involving Condorcet and Borda, and later Lewis Carroll, Hare, Nanson, among others, and culminating in Duncan Black. Ordinalist welfare economics intersected with voting theory, and from Arrow's achievement social choice theory was born.

Social choice theory spans at least three disciplines. Arrow's theorem and much of the discussion of it is motivated by the concerns of welfare economics, that is, what advice should be given to an imaginary social planner who has the task of providing the greatest social welfare to a society (the discussion includes those economists who dispute the legitimacy of such a social-welfare objective). There is also, of course, a large philosophical literature on theories of justice, each of which aspires to provide thorough and coherent arguments about the best way to organize the basic institutions of a society. The concerns of the justice theories overlap somewhat with the concerns of welfare economics, but also typically assert standards of the public good (or some similar objective, such as justice) that are partly or wholly independent of individuals' ordinary rankings of social states. Welfare economics purports to be merely descriptive (or at most to provide hypothetical advice), but theories of justice are frankly normative. Finally, political scientists interested in the formal and empirical exploration of voting and of other political institutions adapted the findings of social choice theory to their purposes. Much of the content of American political science in the last twenty years has been an elaboration of the analogy of political choice to consumer choice first brought to prominence by Arrow and Buchanan in the 1950s. Immediately upon publication of Arrow's *Social Choice and Individual Values*, Little (1952) objected that Arrow's scheme was excessively general in

lumping the economist's interest in social welfare with the political theorist's interest in voting, and Sen (1982, 158–200) later elaborated on the point. My emphasis in this volume is on voting, and I will mostly ignore the concerns of welfare economics and of theories of justice, not from lack of interest, but rather because it would take us too far astray even to state the many issues at stake.

Immediately, democratic voting is different from social-welfare calculations in at least one important respect: it is widely although not universally accepted that citizens should be treated as political equals, in terms of voting that each relevant person should have an equal vote. The economist and the philosopher are free to muse, for example, that one person might be a hundred times better than another in converting life's experiences into some form of satisfaction, so that in the pursuit of equality of welfare such a person deserves fewer resources or perhaps less political influence over the distribution of resources. Or one or the other of them also might propose that the demonstrably competent be granted more than an equal vote or the overly privileged less than an equal vote. The democratic theorist, however, is entitled to the working assumption of formal equality, one vote per one person. The economist or the philosopher might consider that assertion of political equality a defect of democracy as it is presently understood. That would definitely be a minority view, however, and even if it were a correct view I cannot imagine in today's circumstances how the supposed defect would be remedied in practice: if one departure from formal political equality is granted, that only increases the demand that another be granted, escalating into a chaos of exceptions only resolved by a return to formal equality.

Most practical voting systems in use are ordinal: generally, voters are asked to rank-order two or more alternatives and are not asked to express by voting intensity of preferences over alternatives, for example, that Rome is three times as good as Santa Fe and Santa Fe is twice as good as Gary. A cardinal voting scheme, the argument goes, may be vulnerable to misrepresentation: it is in each voter's interest to exaggerate the intensity of her preferences for her favorite choices. Intensity of preference *is* expressed in informal discussions and in democratic debate. The fact that Susie hates seafood because she has a life-threatening anaphylactic reaction to shellfish is enough to overrule a tepid majority's preference for the seafood restaurant – unless Susie insincerely has such a story for every occasion. Susie's claim also relies on some kind of comparability: her life-threatening reaction is much more important than Mark's mild distaste for pizza. A minority's demand to be free from arbitrary deprivation of life and liberty can persuade a majority to desist, perhaps because the majority is well-motivated or perhaps because

the minority threatens to withdraw social cooperation. One member of Congress can tell another that she really needs this vote for her district and would gladly trade votes with others on issues she is nearly indifferent about; there is an element of comparability in that each legislator is allocated only one vote per question. There are voting schemes and practices that under one interpretation approximate cardinality but are not vulnerable to exaggeration: the Borda count (the Borda count need not be justified, however, as an approximation of cardinality), cumulative voting (the voter is allocated a fixed number of points to distribute across alternatives), and under ordinalist majority rule vote trading across a series of issues. Because most voting schemes are ordinal in character, Arrow's theorem and its many offspring are relevant to questions about the comparative desirability of alternative voting schemes and broader questions of institutional design. Any voting scheme assumes some kind of comparability: allocating the same voting power to each person as I have advocated, or weighting votes so as to favor one voter over another, justified by either welfare or nonwelfare considerations. Arrow's theorem assumes ordinal and noncomparable preferences, and voting tends to be ordinal but imposes one or another conception of comparability.

Arrow theorem

The Condorcet paradox of voting, recall, arises from a possible distribution of preference orders among the population such that the aggregate majority vote is a cycle $A > B > C > A$. Arrow's possibility theorem can be understood as a generalization of Condorcet's paradox, applying not only to simple majority voting but also to any social-welfare function that aggregates individual orderings of more than one person over more than two alternative social states. The theorem shows the joint inconsistency of several innocuous-sounding conditions on the social-welfare function. There are many ways to state, informally and formally, the conditions and the results. A good way to begin is with Arrow's (1973) own informal summary of his theorem:

> I stated formally a set of apparently reasonable criteria for social choice and demonstrated that they were mutually inconsistent... The conditions on the social decision procedure follow: (1) for any possible set of individual preference orderings, there should be defined a social preference ordering (connected and transitive) which governs social choices; (2) if everybody prefers alternative A to alternative B, then society must have the same preference (Pareto optimality); (3) the social choice made from any set of available alternatives should depend only on the orderings of individuals with respect to those alternatives; (4) the social decision procedure should not be dictatorial, in the sense that there is one whose preferences prevail regardless of the preferences of all others... The inconsistency of these conditions is in fact a generalized form of the paradox of voting...
>
> As in the original Condorcet case of simple majority voting, all that is meant by the paradox is that it *could* arise for certain sets of individual preference orderings. If individual preference orderings were restricted... then majority voting and many other methods would satisfy conditions (2) to (4).

I shall now state some preliminary definitions, and then the assumptions of the Arrow theorem, along with some interpretations of the formalisms. To begin with, there are alternative social states. A social state, according to Arrow (1963/1951, 17), is a complete description of the amount of each consumption commodity, of labor, of productive resources allocated in the economy, and amounts of all collective activities, ranging from municipal services, to diplomacy, war, "and the erection of statues to famous men." There is an *environment* X of all alternatives, and a set S that is a subset of X. Arrow's framework is quite general, and X could be all possible social states and S all feasible social states. In a specific application such as a presidential election X might be all possible presidential candidates and S all actual presidential candidates. Each individual has a preference *ordering* over all possible social states, and an individual's preferences need not be egoistically oriented, according to Arrow. A *weak* ordering, R, is a generalization of the concept applied to real numbers of "greater than or equal to"; a *strong* ordering, P, is a generalization of "greater than." Strong ordering can be defined in terms of weak ordering: $x\,P\,y$ is defined as $x\,R\,y$ and not $y\,R\,x$. In other words, to say that Italy is better than England is the same as to affirm that Italy is at least as good as England and to deny that England is at least as good as Italy. Indifference can also be defined in terms of weak ordering: $x\,I\,y$ is defined to be $x\,R\,y$ and $y\,R\,x$. To say that Coke is as good as Pepsi is the same as saying at the same time that Coke is at least as good as Pepsi and Pepsi is at least as good as Coke.

Arrow assumes that individuals have consistent preferences over all possible states of the world. A weak preference ordering is *reflexive*: $x\,R\,x$ means that x is at least as good as itself. A weak ordering is also connected, or *complete*: for all x and y, either $x\,R\,y$ or $y\,R\,x$. Someone might be indifferent between Coke and Pepsi, yet she could compare the two. A person who had no weak preference (or strong preference or indifference) between joining the first space voyage to another inhabited planet and spending the same number of years with Socrates would have preferences that were not complete and thus not an ordering. A weak ordering is also transitive: for all x, y, and z, if $x\,R\,y$ and $y\,R\,z$, then

$x\ R\ z$. Arrow's individual would never, for example, prefer Tocqueville to Marx to Mill to Tocqueville. The possibility theorem shows that, if we accept Arrow's conditions, such individual orderings (by definition reflexive, transitive, and complete) cannot be amalgamated into a collective ordering (also by definition reflexive, transitive, and complete). For example, we have seen with the Condorcet paradox that transitive individual preferences can result in a social preference that is intransitive and thus not an ordering. Another voting rule could have the defect that it is incomplete; for example, a voting rule that required unanimity for every decision would in the abstract, if there were any disagreements, result in an incomplete social preference, and would be unable to report social preference or even indifference between some number of alternatives (because of this incompleteness, unanimity rules in practice favor the status quo). To continue with notation, *individuals'* orderings are denoted $R_1, \ldots,$ *Ri* and a *collective* ordering is denoted R (without any subscript).

Now I turn to Sen's widely used formulations.

- "An element x in S is a best element of S with respect to a binary relation R if and only if $\forall y$: $(y \in S: \rightarrow x\ R\ y)$. The set of best elements in S is called its *choice set* and is denoted $C(S, R)$" (Sen 1970, 10). To say that x is a best choice (in the set S with respect to the relation R) means that for all y, if y belongs to S then x is weakly preferred to y. The choice set contains the best elements. We might call them the winners of the contest. If collective preference cycles among the top alternatives, then there is no best element and the choice set is empty.
- "A *collective choice rule* is a functional relation f such that for any set of n individual orderings $R_1, \ldots, R_n$ (one ordering for each individual), one and only one social preference relation R is determined, $R = f(R_1, \ldots, R_n)$." (Sen 1970, 28). For Sen, the collective choice rule is the more general case in which the social preference relation need not be an ordering; it is also made clear that a unique social preference relation is required. Now follows the Arrow theorem, which assumes rather a social preference relation that *is* an ordering. The conditions are labeled O, U, P, I, and D.
- "A *social welfare function* (henceforth, SWF) is a collective choice rule f, the range of which is restricted to the orderings over X. This restriction is to be called condition O on f" (Sen 1970, 41). The collective ranking of alternatives generated by the social-welfare function should be as a collective choice rule unique, and as a social welfare function both complete and transitive.
- "*Condition U (unrestricted domain):* The domain of the rule f must include all logically possible combinations of individual orderings" (Sen 1970, 41). The social-welfare function should accept as input any and all possible individual preference orderings.
- "*Condition P (Pareto principle):* For any pair x, y in X, $[\forall i: x\ P_i\ y] \rightarrow x\ P\ y$" (Sen 1970, 41). For example, if every individual prefers Metallica to AC/DC, then Metallica is preferred to AC/DC in the social preference order.
- "*Condition I (independence of irrelevant alternatives):* Let R and R' be the social binary relations determined by f corresponding respectively to two sets of individual preferences, $(R_1, \ldots, R_n)$ and $(R'_1, \ldots, R'_n)$. If for all pairs of alternatives, x, y in a subset S of X, $x\ R_i\ y \leftrightarrow x\ R'_i\ y$, for all i, then $C(S, R)$ and $C(S, R')$ are the same." Condition I is the condition most difficult to understand and the most frequently misunderstood. The social preference over any given pair of alternatives depends only on individuals' preferences over the same given pair of alternatives; and if individuals' preferences about some third alternative should change, that would not change the social preference over the given pair of alternatives.
- "*Condition D (nondictatorship):* There is no individual i such that for every element in the domain of rule f, $\forall x, y \in X$: $x\ P_i\ y \rightarrow x\ P\ y$" (Sen 1970, 42). If the social preference over every two alternatives is the same as one particular individual's preference over every two alternatives, regardless of the preferences of other individuals, we could call that a dictatorship.
- *Arrow's General Possibility Theorem:* There is no social welfare function (an ordering) that satisfies the conditions of universal domain, Pareto principle, independence of irrelevant alternatives, and nondictatorship.

Here is more on the independence condition. Suppose that we have two individuals, Napoleon who ranks $b > a > c > d > e$, and Josephine who ranks $a > b > c > d > e$. In order to rank a and b, what information does the social-welfare function that obeys Condition I take into consideration? Only that Napoleon ranks b over a and that Josephine ranks a over b, and nothing else. Intuitively, if this were the only information we had, and if we regard Napoleon and Josephine as equals, we would probably conclude that given those individual rankings the social choice between a and b should be a tie. Now suppose that as before Napoleon ranks $b > a > c > d > e$, but that Josephine is different and ranks $a > c > d > e > b$, for Josephine alternative b has dropped from second place to fifth place. What information does Condition I permit to be taken into consideration? Again, only that Napoleon ranks b over a and that Josephine ranks a over b, and by simple majority rule a would tie with b. The alternatives deemed irrelevant in this illustration are any other than a and b. Someone might object: alternative b went from second

to fifth in Josephine's preferences, it seems that Josephine really hates *b*, so it makes more sense now to say that *a* beats *b* rather than that *a* ties *b*. Condition *I* forbids that objection: the social-welfare function accepts only pairwise information from individual preference orders. My explanation of Condition *I* is not typical in the literature, which, if it offers any explanation at all, usually provides an example that makes violation of Condition *I* look silly.

A good proof of the theorem can be found in Sen (1970, 41–46). Naturally, the proof logically depends on the assumed conditions. The work the universal domain condition, *U*, does in the proof is that by allowing any logically possible combination of individual orderings it allows as one possible instance a cyclical profile of individual orderings such as that which gives rise to the Condorcet paradox. If the cyclical profile is excluded for one reason or another, the proof does not go through, as Arrow himself notes in his informal remarks I quoted above. The work that the independence of irrelevant alternatives condition, *I*, does in the proof is to exclude information other than individuals' preferences over pairs. If we could include information about individuals' relative rankings over more than two alternatives, also the proof would not go through. Condition *U* excludes Condorcet voting. Condition *I* excludes Borda and many other methods. Condition *D* excludes the only remaining voting rule, the dictatorship of one.

Just as the Condorcet paradox is used to shock at the elementary level of study, the Arrow possibility theorem is used to shock at a more advanced level. There are thousands of articles varying, extending, and elaborating the theorem. It is often said that the Arrow findings are robust to several variations in the assumptions, and that if the spirit of the conditions is accepted, then there is no magic bullet that puts the problems to rest. That is not quite right. Sen, the social choice theorist who later won the Nobel Prize (a criterion cited by the speakers in my hall of quotations), says that Arrow's conditions are not inescapable commandments. The issue is not the absence of rationally defendable social decision procedures, but rather the importance of disparate conditions that pull in different directions as we evaluate diverse procedures, he says. "We are not at the edge of a precipice, trying to determine whether it is at all 'possible' for us to hang on" (Sen 1995, 11). The usual attention-getting way of stating the result is that the only social-welfare function that satisfies some simple and apparently fair conditions is a dictatorship. A more boring but in my view more appropriate way of stating the Arrow result is to say that if we are required to consider only individuals' rankings of pairs when the collective choice is over more than two alternatives (*I*), or if we must assume that there is no correlation among different individuals' preference orderings (*U*), then there *may* be a cycle in the social choice.

Before we interrogate justifications for the several Arrow conditions, we shall examine the crucial issue of empirical relevance. Throughout his career Arrow has asserted the empirical relevance of cycling, and he came to rely on Riker's findings in support of that belief. In his first major work, *Social Choice and Individual Values* (1963/1951, 3), he introduces the Condorcet paradox and asserts in a footnote that there were cycles in recent Congresses over no federal aid to education, federal aid only to public schools, and federal aid to both public and religious schools, but he does not develop or defend the assertion. In his 1963 (93) postscript to the 1951 volume, Arrow cites Riker (1961) as "the most complete and up-to-date summary of the problem of aggregation of individual choices into collective ones, with particular emphasis on political aspects." Arrow (1963/1951, 120) repeats: "That an intransitive social choice mechanism may as a matter of observed fact produce decisions that are clearly unsatisfactory has been brought out... by Riker... Riker's emphasis is on the possibility that legislative rules may lead to choice of a proposal opposed by a majority." The work to which Arrow refers (Riker 1961) was written while Riker was visiting the Center for Advanced Study at Stanford, where Arrow was in the Economics Department, and Riker acknowledges the generous help and commentary of Arrow on the paper. The paper is a bibliographic survey, and concludes with a declaration of the Rikerian doctrine that many observed majorities are merely apparent because of underlying cycles, providing as examples Riker (1958) on a Congressional appropriations vote (which I show below to be mistaken) and a brief mention of Senate deliberations on the 17th Amendment (later developed in Riker 1982, and which I also show below to be mistaken).

In the same passage, Arrow refers to a few pages in Dahl's (1956, 39–42) *A Preface to Democratic Theory*. Dahl maintains that when roughly equal factions in a group favor mutually exclusive alternatives then democratic rule may be endangered. If democratic procedures do not provide a unique outcome, then the factions may pursue their goals by extra-democratic means, that is, by violence. Deadlock leading to violence would be avoided only if the groups valued maintenance of democratic rule more than they did the nondemocratic pursuit of their goals, according to Dahl. Indeed, "the closer a group approached to an equal division the less valid the majority principle becomes" Dahl believes (1956, 41). Although social choice theorists call majority rule "decisive," that is only by means of the definitional fiat that a tie counts as a decision, in the real world ties are useless when collective action is urgently required.

Dahl's concern is illustrated by the confusion and anguish of the 2000 US presidential election; in the end it mattered less whether Gore or Bush won than that one of them should win (it had to be one of either Gore or Bush; in the circumstances selection of some third candidate would have been truly arbitrary). In a footnote, Dahl includes within his concern a cyclical collective outcome arising from three equally sized groups with a cyclical profile of individual preferences. In other words, for Dahl, a cycle over more than two alternatives is a defect of the same type as a tie between two alternatives. The defect is easily remedied. First, as Dahl mentions, if the population favors democratic rule generally over the nondemocratic pursuit of goals in the particular circumstances of tied outcomes, then democratic stability follows. Second, if ties are a problem then a tie-breaking procedure should be institutionalized in advance: bias to the status quo alternative, or perhaps better a random procedure, or a Republican-majority US Supreme Court, or a neutral constitutional monarch. Notice, however, that the problem of deadlock is not confined to tie votes in a majority-rule setting. If the constitution is of the presidential or Madisonian or Rikerian "liberal" variety, which in the name of checks and balances allows each of many different minorities a veto power over changes from the status quo, deadlocks will be more frequent and thus more dangerous, especially when the status quo is worsened by unanticipated exogenous factors, than if the constitution is of the parliamentary or majoritarian or Rikerian "populist" variety (Stepan and Skach 1994).

Arrow (1960) observes that means of minimizing the deadlocks that arise from the paradoxes of collective choice have evolved in all democratic systems. These brief notes are interesting, because elsewhere Arrow seldom engages in empirical reflections. He rehearses the defects of plurality rule, including that the rule strategically elicits two alternatives and thus disguises possible cycles. Plurality runoff, he says, might exclude a Condorcet winner. Possible ties in US presidential elections are decided by the election going to the US House of Representatives – the theme, apparently following Dahl, is avoidance of deadlock. Finally, in a recent volume Arrow (1997, 5) says, "That there is nothing unlikely about [the Condorcet] paradox has been empirically documented by a number of political scientists beginning with Riker (1958)." It was suggested to me that Arrow's theorem was not intended by its author, nor understood by its audience, to be of empirical relevance; rather it is merely a logical exercise that illustrates a limit case (which is how it should be understood, I believe). I have shown that Arrow himself was motivated by the empirical relevance of his theorem; and that no small part of his audience

shares that motivation is demonstrated by the literature of Riker and his followers. Furthermore, consider Arrow's (1963/1951, 21) methodological credo:

> the present author regards economics as an attempt to discover uniformities in a certain part of reality and not as the drawing of logical consequences from a certain set of assumptions regardless of their relevance to actuality. Simplified theory-building is an absolute necessity for empirical analyses; but it is a means, not an end.

I read that to mean that the model and its logic is of interest only to the extent of its empirical applicability.

The later Arrow (1997, 4) asks, regarding social choice theory, "Does it say that democracy is impossible?" He answers that, "Social choice theory offers only a limited criticism of democratic procedures" (5). He says that although failure to satisfy the theorem's conditions is a legitimate criticism of a procedure, since the failure is universal the theorem alone offers no basis for differentially evaluating alternative social choice mechanisms (8; including, I would add, the market). Instead, "In the case of real social choice procedures, we have to consider the frequency with which intransitivities [and other violations] occur. This is not the sort of result I like, but that is the way the world is" (8).

Arrow implicitly rejects the claim by Riker and his followers that the general possibility theorem renders democracy impossible. Arrow explicitly accepts the claim by Riker and his followers that empirical cycles have been robustly demonstrated, and he has always held that the relevance of his theorem is an empirical question. What is the frequency of social intransitivities?

> [E]mpirical observations of a wide variety of actual collective decision-making processes indicate that cyclical majorities are very rare. Thus, cycles do not appear to be a real problem for group decision-making although some paradoxes may occur which may go undetected. (Feld and Grofman 1986)

Actual observations about majority cycling are scarce, because elections rarely generate data on pairwise comparison among all alternatives or a ranking of all alternatives (Gehrlein 1983). In the limited instances where such data were available from experimental subjects or by inference from a sample of actual legislative situations, the absence of a Condorcet winner tended to be infrequent, according to the summary by Gehrlein (1983). Since then, whenever good data were available and analyzed, cycles have been shown to be infrequent, as we shall now see.

What's the frequency?

Dobra (1983) solicited real election data from the readers of the journal *Public Choice*. This is no random sample: the readers of *Public Choice* are motivated to notice and report cycles. The cases reported were mostly faculty searches and from a small experiment, with small numbers of voters (median 10, range 4 to 27) and larger numbers of alternatives (median 5, range 3 to 37). There were three cycles including ties (e.g., $A > B > C \sim A$) and one cycle not including a tie out of the 32 cases. Although cycles were infrequent, Dobra holds that the infrequency does "not repudiate the work of the disequilibrium theorists" (247).

Recall the study by Chamberlin, Cohen, and Coombs (1984) of five different presidential elections of the American Psychological Association (APA) where voters rank-ordered all candidates, permitting hypothetical comparison of voting rules. The data also permit examination of aggregated preference orders for the presence or absence of cycles. Only about half the voters in these elections ranked all five candidates; the authors generated complete ballots in one condition by filling in remaining preferences *randomly* (as if voters were indifferent to unranked candidates) and in another condition by filling in remaining conditions *proportionately* (as if voters were uninformed about unranked candidates). Either way, there was a transitive majority ordering of the candidates in all five elections. The actual APA preference orders yield zero cycles. With the same numbers of alternatives and voters as in the APA election, but rather an impartial culture, there would be a 24 percent expectation of cycles. There are no cycles in the APA elections because preference orders are not randomly distributed as they are under the impartial-culture assumption. Natural preference orders have some minimum of structure, such as to preclude cycles nearly all of the time. As mentioned above, alternative voting rules picked the Condorcet winner about 80 percent of the time, and the second-ranked Condorcet candidate the rest of the time, again because of minimal similarity among preference orders.

Niemi and Wright (1987) looked at thermometer ratings (respondents rate 0 to 100 when 0 means very unfavorable and 100 very favorable) of 14 politicians who were potential or actual candidates for the 1980 American presidential election, from a nationally representative sample of US voters. If *all* preference orders within a group are *single-peaked* (to be detailed below) and thus unidimensional then no cycle is possible. If only one of the voter's rankings within a group fails the single-peakedness criterion then a cycle is possible but it is most improbable; and the higher the proportion of single-peaked preference orders within the group the less likely is a cycle (Niemi 1969). For three-, four- and five-candidate

groups, Niemi and Wright found that the observed proportions of unidimensional preference orders in each condition are much higher than what would be expected by chance under the impartial-culture assumption; and those observed proportions were such that the probabilities of cycles were 0.04 or less. Further, the ten three-candidate rankings with the *worst* unidimensionality were sampled 1,000 times each; there were 29 cycles in these 1,000 samples, and 22 of those cycles involved the most obscure figure of the 14 candidates (Lucey, John Anderson's vice-presidential running mate). For the *least* unidimensional four-candidate rankings, top cycles were found in 154 out of 1,000 samples, and 137 of those occurred in two of the ten rankings. They also found absence of a relationship between unidimensionality on the one hand and on the other hand more distinguishable candidates, voter education, voter partisanship, judging better-known candidates, longer campaign exposure, or within-party judgments. Curiously, however, the dimension used by voters in their data did not appear to be the standard left–right dimension, but rather perhaps one of likability.

Some very good data from candidate elections in private organizations in Great Britain were analyzed by Feld and Grofman (1992). These 36 elections, with between 3 and 29 candidates and between 9 and 3,422 voters, were conducted by the single transferable vote procedure, which requires that voters rank-order candidates. The authors state that although strategic voting is possible in principle under single-transferable vote, the necessary calculations are too complex for it to occur in practice (Bartholdi and Orlin 1991 show that it is NP-complete, that is, too computationally complex, except in special cases, for voters to be strategic under single transferable vote). From the data of the voters' rank-orderings, pairwise comparisons can be constructed. Every one of the reconstructed elections had a Condorcet winner. In 34 of the 36 elections the Condorcet winner and the Borda winner coincided (so the results speak as well against Riker's claim that democracy is inaccurate). Only 0.5 (one-half of one) percent of the linearly ordered triples in the sample universe were cyclic; 24 of the 36 elections had no cycles whatsoever, the largest percentage of cycles in any election was 2.0 percent; and almost all cycles were among alternatives adjacent in Borda scores (meaning that they were among close alternatives). Since many observers believe actual cycles are infrequent, with those asserting otherwise emphasizing the absence of data to make a confident determination, I agree with Feld and Grofman that their results are important.

Felsenthal, Maoz, and Rapoport (1993) review a mostly overlapping set of elections, and make similar findings. They observe that among elections with eight or fewer candidates there is only one instance of a (minor)

cycle, and among elections with more than eight candidates there are 13 out of 19 instances of (minor) cycles. Felsenthal and Machover (1995) consider an expanded set of 92 elections, and replicate the observation: 7.5 percent of elections with eight or fewer candidates contain cycles, but 56.4 percent of elections with more than eight candidates do, and they remark on the sharp jump in occurrence of cycles beyond eight candidates. I shall propose an explanation for this in the next chapter. Of the 92 elections, 26 contained cycles; no cycle involved all candidates, and only two were cycles at the top of the preference order. Of the 26 elections with cycles, winners were immediately obvious in 19 of the elections, either because the cycle was not at the top, the number of slots to fill exceeded the length of the cycle, or indifference relations within cycles appropriately indicated winners.

Radcliff (1993) reports high unidimensionality derived from studies of American presidential elections from 1972 to 1984: 77 percent to 85 percent single-peakedness in years with three major candidates in the primary and general elections (at 75 percent unidimensionality the expectation of a cycle is less than 1 percent, and at 80 percent the expectation is almost zero, Radcliff 1994); and 50 percent in 1980 when there were five major candidates. Radcliff (1993) also finds that most *individual* voters' rankings are transitive but that individual intransitivities increase with number of candidates. If intransitive voters are removed from consideration in the five-candidate case, single-peakedness goes up to 70 percent. Radcliff (1994) uses data from the same election studies to examine the transitivity of *collective* rankings. There was a Condorcet ordering in each American presidential election studied (that is, no cycles); the ordering corresponded to the standard left–right dimension of understanding; and each actual election picked the Condorcet winner. Thus: 1972, Nixon > Humphrey > McGovern; 1976, Carter > Ford > Reagan; 1980, Reagan > (Carter ~ Bush) > Anderson > Kennedy; 1984, Reagan > Hart > Mondale.

Van Deemen and Vergunst (1998) continue the elusive quest for the empirical cycle. From the Dutch parliamentary election studies of 1982 (13 parties), 1986 (12 parties), 1989 (9 parties), and 1994 (9 parties) they have survey data on respondents' preferences over the alternatives. If preference rankings were random, as under the impartial-culture assumption, then there would be about a 50 percent chance of cycles. There are, however, *no cycles in their data, not anywhere in the rankings, not in any of the elections*. The authors find the results surprising: "for some reason or another cycles in large elections are scarce" (485). Kurrild-Klitgaard (2001a) investigates cycles in Danish national election surveys. Respondents ranked 11 parties in 1973, 9 parties in 1994, 11 parties in 1998;

the policies of the 8 parliamentary parties in 1994; 10 party leaders in 1973, 8 in 1994, and 10 in 1998; the importance of 28 issues in Danish politics in 1987/1988, and of 16 issues in 1994; 4 important goals in 1994; 12 public budget alternatives in 1990, and 20 such alternatives in 1994. With the number of respondents and often large number of alternatives, the prediction from simulations based on impartial-culture assumption would be a large proportion of cycles. For example, with an impartial culture, many voters and 20 alternatives, the prediction is a 68 percent incidence of cycles (Gehrlein 1997, 179). There was, however, *only one trivial cycle in this entire set.* For the 20 budget alternatives in 1994 there was a Condorcet ranking for the first 14 alternatives, and then a cycle over what would be alternatives 15 through 18, and then transitive order again over alternatives 19 or 20. Otherwise, no cycles.

Finally, Regenwetter and Grofman (1998) used a probabilistic method to estimate rankings from data in seven out of ten real approval-vote elections, and found only a small chance of a cycle in one of the elections. We should remember the bias against publishing negative findings. No doubt many people over the last thirty years have thought that it would be intellectually and professionally satisfying to demonstrate a real instance of cycling, yet the positive claims of cycling we have from the entire political universe can be counted on one's fingers and toes (and, as we shall see, even these claims collapse under scrutiny). Where is the pervasive political disequilibrium? Shepsle and Bonchek (1997, 50–51) ask:

> Is [group intransitivity] merely an arcane logical possibility, a trick foisted on the unknowing student by professors, philosophers, and textbook writers? Or is it a profound discovery, the stuff from which important insights about political philosophy and social life are made. In our opinion, the answer lies much closer to the latter.

In my opinion, the answer is closer to the former rather than to the latter.

Riker (1965, 52) warned that the Arrow theorem is no mere "mathematical trick without practical significance" and set out to show that the paradox of voting does occur and is of tremendous importance in committees and legislatures. At first, he estimated that whenever on important issues a proposal loses to the status quo, half the time it is due to a manipulated cycle. Later, Riker estimated that cycles afflict 10 percent of legislative votes (Bell 1974, 308). Still later, Riker (1982, 122–123) acknowledged there is a tendency to similarity among preference orders that reduces the likelihood of cycles: "there is good reason to believe that debate and discussion do lead to . . . fundamental similarities

of judgment." However, the possibility of manipulation increases the likelihood of cycles, he argues. The net result: there are few cycles on unimportant issues, but more cycles the more important the issue is to the manipulators. More precisely:

> quite a wide variety of rather mild agreement about the issue dimension guarantees a Condorcet winner . . . not all voters need display the agreement to obtain the guarantee . . . agreement about dimensions renders uncontrived cyclical outcomes quite rare . . . intransitivities only occasionally render decision by majoritarian decisions meaningless . . . at least when the subjects for political decision are *not* politically important. When, on the other hand, subjects are politically important enough to justify the energy and expense of contriving cycles, Arrow's result is of great practical significance . . . on the very most important subjects, cycles may render social outcomes meaningless. (Riker 1982, 128)

I will argue in Chapters 9 through 17 that Riker is unable to demonstrate the existence of a cycle on any issue, minor or major.

Riker later (1990b, 179) granted that, "Poole and Rosenthal . . . have shown with large empirical studies of congressional voting that, in the absence of grand manipulation, a considerable part of political life is unidimensional." Poole and Rosenthal (1997) dedicate their book to Riker, their "teacher, friend, and colleague." Their spatial analysis of *all* roll-call votes in the 1st through 100th Congresses of the United States shows that about 85 percent of all votes can be accounted for in two dimensions. Moreover, "Except for two periods of American history, when race was prominent on the agenda, whenever voting could be captured by the spatial model, a one-dimensional model does all the work" (227). The first and overwhelmingly important dimension is what we popularly understand as the standard left–right dimension. The second dimension explains only about 2 percent of the 85 percent captured. The second dimension varies from Congress to Congress, and varies from public works to currency to tariffs and other issues; but was most salient as slavery in the period before the Civil War (the 37th Congress in 1850 most poorly fit the spatial model) and as race relations in the civil rights era of the 1950s and 60s. Testing for third, fourth, and greater dimensions on the whole does not explain meaningfully more than the two-dimensional model. Since the mid-1970s, the Congress has become increasingly and is now almost wholly unidimensional (Poole and Rosenthal 1999). A unidimensional issue space implies no cycles; and the mostly unidimensional issue space discovered by Poole and Rosenthal implies very few cycles.

Originally, I suspected that this unidimensionality was somehow a product of the American two-party system and thus not evidence for

a strong tendency to unidimensionality in politics. The Poole–Rosenthal (1999) methods, however, have recently been applied to votes in the European Parliament (1989–1997), the British Parliament (1841), the French National Assembly (1951–1956), the Czech Parliament (1993–1997), the Polish Parliament (1995), and the United Nations General Assembly (1946–1996). The percentage of votes correctly classified by a single dimension of analysis ranges from 85.9 percent in the UN (1954–1969) to 94.2 percent in the Czech Parliament. The Czech Parliament is a multiparty system, and the United Nations comprises the diverse interests of six billion people. It is possible that the apparent unidimensionality is an artifact of the Poole–Rosenthal methodology.

Budge (1993) and coworkers examined all party manifestoes or platforms from 1945 to 1981 in 23 democratic countries and applied factor analysis. They found that one dimension, the standard left–right dimension, best explained the data. After reporting the findings of Poole and Rosenthal and of Budge, Riker (1993, 4) acknowledged that "issue spaces tend to be one-dimensional over time." He responded that second dimensions would be of relevance, presumably with respect to manipulation, in the short run.

On his own terms, Riker's claim of meaninglessness now stands only on incidents of grand manipulation. Even if, as Riker maintains, all cyclical manipulations are difficult to detect, he must be able to demonstrate some instances from the rich universe of politics, especially since cycles are supposed to be associated with the most important issues on which we would have the most information. Otherwise, his claim would have nothing but the glory and the shame of an untestable empirical claim. If there are only a handful of such incidents, then the meaninglessness claim fails, and it utterly fails if the handful do not withstand scrutiny. Later we shall investigate in detail the topics of strategic voting and agenda manipulation; for now accept that an instance of legislation defeated by a killer amendment implies the presence of natural cycles or of the manipulatively contrived cycles that Riker stresses. With respect to strategic voting, Poole and Rosenthal (1997, 147) "found very few bothersome needles in our haystack of 37,000 roll calls." The three instances of successful killer amendments identified in the political science literature are the three recited by Riker in *Liberalism against Populism*, known as the Wilmot Proviso, the Depew–Sutherland amendment, and the Powell amendment, and each does have to do with race, often the second dimension in American politics, according to Poole and Rosenthal (1997, 162). With almost fifty years of controversy, and strong professional incentives for unveiling grand manipulations, these are the three that we have from the universe of American Congressional roll-call votes to support the

proposition that democracy is meaningless. In later chapters, I will show that Riker's accounts of these three events are mistaken. The jewel in the crown of Riker's examples of grand manipulation (not related to Congressional roll calls) is his famous allegation of a cycle in the 1860 American presidential race that resulted in the arbitrary election of Lincoln and consequently the Civil War. As promised, I will later show that his account of that election is clearly mistaken and that there was no such cycle.

In a review article on the subject, Enelow (1997) writes that, "cycling and majority rule is one of the most heavily researched areas of public choice" (149), yet, citing only Riker, he acknowledges that "basically, the empirical literature testing the theory we have described consists of a small set of examples" (160). If I succeed in showing that Riker's and others' examples are mistaken, then the cycling hypothesis must die from lack of evidence. Further, it would be shown that the Arrow theorem cannot be interpreted to conclude that democracy is meaningless.

Justifying the theorem's conditions

The definition of a social-welfare function requires that both individual and social preferences be orderings – complete and transitive ordinal rankings. Condition *I* goes beyond requiring individual orderings, and in addition requires that all voting rules proceed by pairwise comparison. These issues will be discussed in depth in Chapter 6.

Condition *P* (if all voters prefer Metallica, then society prefers Metallica) is not as innocent as it first looks. That's because the Arrow theorem is not an engineering guidebook, but rather it is a logical exercise: it need not be that any of the conditions are actually violated, rather the transgression occurs if one of them *could* be violated. Condition *P* rules out such social-choice rules as: do whatever the Bible says to do. It may be that all voters want society to do what the Bible says to do, but the point is, if the citizens were to change their view about that, the social-choice rule would still dictate doing what the Bible says to do. Consider, among other concepts, Rousseau's concept of the general will, inalienable rights to life and liberty, the US Constitution, or Rawlsian justice. Each violates Condition *P*, to the extent that each is not based on aggregation of individual orderings into a social ordering. It could be that the unanimous will of all (aggregated votes) is identical to the general will (the true or right decision), but it *could* be that everyone is mistakenly opposed to the general will; and because it could happen, the general will as a decision rule is in violation of *P*. It may that everyone agrees that I have an inalienable right to life and liberty, but it is logically possible that everyone, including me, would vote to deny that right. Thus, the right is in violation of *P*.

Condition *D* – no one person determines the social ordering regardless of the preferences of other individuals – is of direct normative relevance to democratic theory. The condition is quite thin, however. Other than Conditions *P* and *D* the Arrow theorem is noncommittal about democracy, as the theorem declares impossible both democratic voting rules that give each citizen one vote and undemocratic voting rules that give more weight to some classes or exclude other classes of persons or establish any dictatorship of two. Notice also that in the absence of some independent justification, Arrow's nondictatorship condition seems to bring a worry relating to the interpersonal comparison of welfare into the scheme, contrary to its logical positivist foundations. If there is no way of comparing the welfare of one person to another, then why should we object if one person gets to decide everything? It could be, and many dictators act as if it were so, that the satisfactions of the dictator are worth 1,000 times everyone else's put together.

That leaves Condition *U*, unrestricted domain, which says that the domain of the social-choice function includes all logically possible individual orderings of the alternative social states. When social preferences cycle, there is no social choice, according to the definitions. Condition *U* contributes to the impossibility result because only cyclical profiles of individual preferences yield cyclical social orderings. If the domain of individual preferences were limited so as to exclude cyclical profiles, then there would be no impossibility result. Two justifications typically are offered for Condition *U*. First, that Arrow's theorem seeks generality, and *U* is the most general assumption. But that is not so. Condition *U* requires individuals to have complete and transitive orderings. Obviously they don't in real life, and a more general condition would be to permit individuals to have some incomplete and some intransitive preferences. The problem with the more general condition is that it would detract from the rhetorical force of the Arrow result, since all that would be shown from the more general condition is that incomplete and intransitive individual preferences may aggregate into incomplete and intransitive social preferences.

Second, it is said that to exclude some preference orderings would be tyrannical. Condition *U* has been interpreted to mean that "citizens should be free to prefer any policy option at all and to rank any options in any way they want, meaning that no institution should have power to declare certain choices out of bounds at the start" (Pildes and Anderson 1990, 2,132). Any domain restriction that would mitigate the Arrow theorem's impossibility result would have harsh consequences for those whose

preference rankings fall outside the domain restriction. We would have to "restrict entry into the community to those having preference orderings that do make collective choices possible" or if it is already too late "they must somehow be isolated and excluded from the community, or an impossibility result can again emerge" (Mueller 1989, 392–393). In the next chapter, I shall argue that if the question is about the public interest, then individual preferences are naturally sufficiently similar to one another's to avoid cycling most of the time. And if the question is fixed-sum redistribution, then destructive self-seeking preferences *should* be excluded from public consideration (as taught in kindergarten).

5 Is democracy meaningless? Arrow's condition of unrestricted domain

Introduction

Given theoretically predicted instability, why the empirically observed stability? There are several types of answers: that stability is an illusion because we are unable to detect the manipulation that occurs (Riker 1982); that stability is due to institutional devices (e.g. Shepsle and Weingast 1984); that such institutional devices are themselves pervasively unstable (Riker's rejoinder 1980a); that stability is due to similarity in preference rankings among the population, and to preferences for fair distribution; or is due to some other defect in the models. The counterempirical outcome of Arrow's theorem puts us on notice that one or another of its conditions must be misconceived.

I begin with similarity among individuals' preference rankings, a challenge to the realism of Arrow's condition of unrestricted domain. The theorem's impossibility result is a logical possibility but not an empirical probability, I shall argue. One kind of similarity in preference rankings is disastrous though: if majority-rule voters divide up a fixed good, and if each is motivated solely by self-interest and not at all by fairness, then we are guaranteed instability. Contrary to theoretical prediction, however, democratic legislators are typically universalistic rather than factional on distributional questions. This may be due to uncertainty about the future, or a direct concern for fairness, or independently motivated reciprocity, or public deliberation, or due to some combination of these devices. Empirical work shows that citizens vote judgments of general welfare rather than personal welfare. An individual voter almost never affects the outcome of an election, hence she is free to express her disinterested sentiments rather than her interests, and this may explain the empirical finding. Further, there is empirical evidence that Americans overestimate the prevalence of self-interest. If there is not sufficient fairness in the population to tame Condorcet-voting instability, I argue, then there are other acceptable voting rules that avoid cycles altogether.