CHAPTER 24

The impossibility of a social ordering

> The only orthodox object of the institution of government is to secure the greatest degree of happiness possible to the general mass of those associated under it.
>
> Thomas Jefferson

> A really scientific method for arriving at the result which is, on the whole, most satisfactory to a body of electors, seems to be still a *desideratum*.
>
> Charles Dodgson
> (Lewis Carroll)

The Bergson-Samuelson SWF has been constructed analogously to the individual's utility function. Just as the individual chooses bundles of commodities to maximize his utility, society must choose an allocation of commodities across individuals to maximize its welfare. That consumers make choices to maximize their utility follows almost tautologically from the definition of rationality. In extending the idea of maximizing an objective function to the level of society, however, more is involved than just rationality. Embedded in the characteristics of the welfare function and the nature of the data fed into it are the value judgments that give the SWF its normative content, as the discussions of Bergson (1938) and Samuelson (1947, ch. 8) make clear.

An alternative way of analyzing individual behavior from assuming that individuals maximize their utility is to assume various postulates about individual rationality that suffice to define a preference ordering, and allow one to predict which bundle an individual will choose from any environment. Again by analogy, one can make various postulates about social decision making and analyze society's decisions in terms of social preference orderings. What choice should a society make from a given environment? Again, however, in shifting from the individual to the societal level, the postulates change from simply defining rationality to expressing the ethical norms of the community. This is important to keep in mind because some of the axioms *sound as if* they simply require collective rationality, and some writers have so interpreted them. This is not the course followed here. In discussing each axiom, we emphasize its normative content.

The first and most important attempt to define an SWF as a social ordering satisfying a few, basic ethical axioms was made by Arrow in 1951 (rev. ed. 1963). Although some of Arrow's discussion of the individual axioms seems to mix ethical and rational considerations, the overriding objective of the inquiry is normative, and

our emphasis on the normative characteristics of the axioms does not seem out of place. Arrow, himself, has accepted the interpretation of these axioms as indicating the basic value judgments to be incorporated in the community's social contract or constitution,[1] and this is perhaps the best way to look at them. The question then is this: What ethical norms are we to impose on the social choice process, and what collective choice processes satisfy these axioms? The answer is disappointing. Given but a few fairly weak and ethically uninspiring axioms, no process (voting, the market, or otherwise) exists that satisfies them.

We begin by briefly stating the axioms and sketching the impossibility proof, after which we turn to a more detailed examination of the axioms.

24.1 Logic of the proof

I follow Vickrey's (1960) restatement of the postulates and proof, since they are simpler and shorter.

1. *Unanimity* (the Pareto postulate): If an individual's preference is unopposed by any contrary preference of any other individual, this preference is preserved in the social ordering.
2. *Nondictatorship*: No individual enjoys a position such that whenever he expresses a preference between any two alternatives and all other individuals express the opposite preference, his preference is always preserved in the social ordering.
3. *Transitivity*: The social welfare function gives a consistent ordering of all feasible alternatives. That is, $(aPbPc) \rightarrow (aPc)$, and $(aIbIc) \rightarrow (aIc)$.
4. *Range* (unrestricted domain): There is some "universal" alternative u such that for every pair of other alternatives x and y and for every individual, each of the six possible strict orderings of u, x, and y is contained in some admissible ranking of all alternatives for the individual.[2]
5. *Independence of irrelevant alternatives*: The social choice between any two alternatives must depend only on the orderings of individuals over these two alternatives, and not on their orderings over other alternatives.[3]

[1] This interpretation was first put forward by Kemp and Asimakopulos (1952) and was subsequently endorsed by Arrow (1963, pp. 104–5).

[2] Arrow's statement of the axiom is as follows:

> Among all the alternatives there is a set S of three alternatives such that, for any set of individual orderings $T_1, \ldots, T_n$ of the alternatives in S, there is an admissible set of individual orderings $R_1, \ldots, R_n$ of all alternatives such that, for each individual i, $x R_i y$ if and only if $x T_i y$ for x and y in S (Arrow, 1963, p. 24).

[3] Vickrey states this postulate somewhat differently, but his proof relies on it in this form. This statement of the axiom also differs from Arrow's original statement of it, and others existing in the literature. Arrow's statement is as follows:

> Let $R_1, \ldots, R_n$ and $R'_1, \ldots, R'_n$ be two sets of individual orderings and let $C(S)$ and $C'(S)$ be the corresponding social choice functions. If, for all individuals i and all x and y in a given environment S, $x R_i y$ if and only if $x R'_i y$, then $C(S)$ and $C'(S)$ are the same (Arrow, 1963, p. 27).

For a statement of the axiom in the present way, and impossibility proofs based on it, see Sen (1970a).

Condition 4 perhaps requires an additional word of explanation. The notion of a universal alternative is not crucial here. What is implied by the range axiom is that the social choice process allows any possible ordering of the three alternatives x, y, and u. The process is not established in such a way as to rule out possible orderings.

The theorem states that no SWF satisfies these five postulates. To understand the significance of the theorem, it is useful to run through the proof, again following Vickrey. We first define a decisive set D.

Definition of decisive set: *A set of individuals D is decisive, for alternatives x and y in a given social welfare function, if the function yields a social preference for x over y, whenever all individuals in D prefer x to y, and all others prefer y to x.*

Proof:

Step	*Justification*
1. Let D be a set of individuals decisive for x and y	Assumption
2. Assume for all members of D $xPyPu$, and for all others (those in C) $yPuPx$	Range
3. For society xPy	Definition of D
4. For society yPu	Unanimity
5. For society xPu	Transitivity
6. But for only members of D is xPu	Assumption
7. Society must prefer x to u regardless of changes in rankings of y or any other alternatives	Independence
8. D is decisive for x and u	Definition
9. D is decisive for all pairs of alternatives	Repetition of steps 2–8
10. D must contain two or more persons	Nondictatorship
11. Divide D into two nonempty subsets A and B	Assumption
12. Assume for A $xPyPu$ for B $yPuPx$ for C $uPxPy$	Range
13. Since for members of A and B, yPu, for society yPu	Definition of D
14. If for society yPx, B is decisive for y and x	Definition of D
15. If for society xPy, then for society xPu	Transitivity
16. But then A is decisive for x and u	Definition of D

In either case, one of the proper subsets of D is decisive for a pair of issues, and therefore by step 9 for all issues. Steps 10–16 can be repeated for this new decisive set and then continued until the decisive set contains but one member, thus contradicting the nondictatorship postulate.[4] □

[4] This literature is replete with this form of "Chinese boxes" proof to uncover the dictator. For an important variant thereon with infinite numbers of voters, see Kirman and Sondermann (1972).

The intuition underlying the proof runs as follows: the unrestricted domain assumption allows any possible constellation of ordinal preferences. When a unanimously preferred alternative does not emerge, some method for choosing among the Pareto-preferred alternatives must be found. The independence assumption restricts attention to the ordinal preferences of individuals for any two issues when deciding those issues. But as we have seen in our discussions of majority rule, it is all too easy to construct rules that yield choices between two alternatives but produce a cycle when three successive pairwise choices are made. The transitivity postulate forces a choice among the three, however. The social choice process is not to be left indecisive (Arrow, 1963, p. 120). However, with the information at hand – that is, individual ordinal rankings of issue pairs – there is no method for making such a choice that is not imposed or dictatorial.

24.2 Relaxing the postulates

To avoid the impossibility result, the postulates must be relaxed. Before doing so, however, let us consider the significance of the theorem as it stands, for its significance stems precisely from the weakness of the postulates as now stated. Although, as we shall see, these axioms are somewhat stronger than they might first appear, they are far weaker than one would wish to impose at the constitutional stage to satisfy reasonable notions of distributional equity. For example, there is nothing in the axioms to preclude one group of individuals, as long as it has more than one member, from tyrannizing over the others, if it stays on the Pareto frontier.[5] Even allowing this and still other violations of our ideas of equity, we cannot find a process to choose from among the Pareto-optimal set that satisfies these axioms.

Space precludes a complete review of all modifications of the postulates that have been made to produce either possibility theorems or new impossibility results. Instead, we focus on modifications of particular relevance to public choice.

Relaxing unanimity and nondictatorship hardly seems worth discussing if the ideals of individualism and citizen sovereignty are to be maintained.[6] These two axioms clearly illustrate that what we are engaged in here is a normative exercise. There is nothing particularly irrational about selecting one individual and allowing him to make all decisions for the community; indeed, arguments for an omniscient dictator have been around at least since Plato's eloquent defense of this alternative in *The Republic*.[7] But such arguments are inconsistent with our most basic democratic ideals. Special mention should also be made of Hobbes's defense of monarchy (1651). To Hobbes, there was one issue on which all preferences were identical: life in anarchy was terrible and inferior to life under a unanimously accepted dictator. If one made the other postulates part of the Hobbesian contract, one might construct a new defense of autocracy; and, of course, in practice the dictatorial solution to the

[5] See Sen's amusing example (1977a, p. 57).

[6] However, see Little (1952).

[7] Bell (1973) presents the modern version of this position. After citing Arrow's proof at a number of places to indicate the difficulty that purely democratic processes have reaching decisions, he opts for choice by technocratic experts who form the ruling elite in the postindustrial society.

uncertainties and deadlocks of social choice is very popular. Empirically, it might be interesting to investigate the frequency with which dictatorial governments replace democratic ones following apparent deadlocks of the latter stemming from voting paradoxes. The other three axioms require more detailed discussion.

24.2.1 *Transitivity*

Arrow's reasons for requiring that the social choice process produce a consistent social ordering appear to be (1) "that some social choice be made from any environment" (1963, p. 118), and (2) that this choice be independent of the path to it (p. 120). These are in fact different requirements, and neither of them requires the full force of transitivity.

The requirement that the social choice process should be able to make some choice from any environment seems the easiest to defend, deadlocks of democracy being an open invitation to dictatorship. But to achieve the goal one does not have to assume the existence of a social preference ordering defined on the basis of all individual preference orderings. To make choices one needs only a *choice function* that allows one to select a best alternative from any set of feasible alternatives (Sen, 1970a, pp. 47–55; Plott, 1971, 1976). Transitivity is not required. Either quasi-transitivity or acyclicity will suffice (Sen, 1970a, pp. 47–55). Both of these conditions are milder than transitivity. Quasi-transitivity requires transitivity of the preference relation, but not of indifference; acyclicity allows x_1 to be only "at least as good as" x_n even though $x_1 P x_2 P x_3, \ldots, x_{n-1} P x_n$. Possibility theorems have been proven by replacing transitivity by either of these and retaining the other Arrow axioms. Gibbard (1969) has shown, however, that requiring a quasi-transitive ordering of the social choice function produces an oligarchy that can impose its unanimous preference on the rest of the community; and Brown (1975) has shown that acyclicity gives veto power to *every* member of a subset of the committee that Brown calls a "collegium."[8] Thus, as one relaxes the consistency requirement from transitivity to quasi-transitivity, and then to acyclicity, dictatorial power becomes spread and transformed, but does not disappear entirely. Requiring the social decision process to be decisive in some sense vests one individual or a subset of individuals with the power to decide, or at least block, any outcome.[9]

Although relaxing the transitivity axiom has some advantage in spreading dictatorial power across a wider group, it incurs the additional cost of introducing a degree of arbitrariness into the process (Sen, 1970a, pp. 47–55). Under quasi-transitivity, for example, aIb and bIc can exist along with aPc. Then, in a choice between a and b, society can pick either, but if c is added to the set, society must pick only a. If a, b, and c are points on the Pareto frontier, there will be distributional consequences to the choice of any one. Those favored under b may question the ethical underpinnings of a process that makes their fate dependent in such a seemingly capricious way on the set of alternatives under consideration.

[8] See also Blau and Deb (1977).

[9] For further discussion of this point, see Brown (1973), Plott (1976, pp. 543–6), and Sen (1977a, pp. 58–63).

The gain from relaxing the transitivity axiom is further reduced when one considers the restrictions that must be placed on the patterns of individual preference orderings to ensure that either quasi-transitivity or acyclicity holds. For majority rule, at least, the conditions that are necessary and sufficient for acyclicity are the same as those required for quasi-transitivity, and these in turn will also ensure transitivity when the number of individuals is odd.[10] Thus, if the property of having a choice made from every environment is to be maintained, there appears to be little lost by sticking to the full transitivity requirement.

The intuition behind requiring the final outcome to be independent of the path to it is somewhat different. Here, to begin with, a *path to* the final outcome is obviously assumed. That is, a choice is not made from the full set of all possible candidates, but instead winners are selected from subsets of the full issue set. These in turn are pitted against one another in some manner, and a given path is followed until a final choice set is found. The requirement that the social choice process should be path independent amounts to the requirement that the final choice set should be independent of how the initial subsets are formed out of the full issue set (Plott, 1973).

Path independence is related to and in fact implies another condition that has received much attention in the literature, Sen's (1969) property α, already introduced in Chapter 23. Property α states that if x is a member of the choice set defined over the full set of alternatives S, then x is a member of the choice set of any proper subset of S of which it is a member. Property α is one of a group of *contraction-consistent* properties that have been investigated.[11] As the set of alternatives is contracted, x must continue to be chosen as long as it is one of the alternatives. The intuitive notion here is perhaps obvious: if x is the best chess player in the world, then he is also the best chess player in London. Path independence in this context requires that x's emergence as champion be independent of how the original runoff matches were ordered. This latter requirement is obviously stronger than the former, which explains why path independence implies the α-property, but not the reverse.

Complementary to α and the other contraction-consistent properties are a set of *expansion-consistent* properties such as the β property (Sen, 1969, 1970a, 1977a). The β property states that if x and y are both members of the choice set for some subset S_1 of the full set S, then x can be a member of the choice set of S if and only if y is. Returning to our chess champion examples, if x and y tie for the chess championship of England, then the β property requires that they both be among those who tie for the chess championship of the world, if either one of them ties for the world championship. As Sen pointed out, it is quite plausible in examples such as these for two individuals to tie in a local contest, but one goes on to beat all others and emerge the world champion. Thus, although β may be a reasonable constraint to place on some choice processes, as when contestants are measured in a single dimension like weight, it does not seem as reasonable when the candidates are measured (or compete) in several dimensions. Since issues arising in a social choice context are likely to take the latter form, it is quite possible that a social

[10] See Sen and Pattanaik (1969), Inada (1970), and Sen (1977a).

[11] See Sen (1977a, pp. 63–71).

decision process would violate property β and still not seem inherently irrational or unfair.

Thus, of the two types of properties, the intuitive support for contraction-consistent or path-independent properties seems much stronger than for expansion-consistent properties of the β type. What we seek is the social choice, or set of choices, that defeats all others. Having found such a choice, it would be comforting to know that its selection was independent of the chance way in which earlier contests were established (path independence), and that it could compete again against any subset of losers and still emerge a winner (α property). Unfortunately, it is path independence and the α properties, even in their weakest forms, that lead to dictatorial or oligarchical social preference orderings; the only possibility theorems that have been proven impose only expansion-consistent properties of the β type.[12]

Let us consider somewhat further what is at stake if we abandon all vestiges of the transitivity axiom. Requiring that the social choice process satisfy this axiom is motivated in part by the desire to avoid the embarrassment of inconsistency and arbitrariness. But this view in turn seems to stem from the belief that, just as it is *irrational* for an individual to exhibit inconsistent preference orderings, it is *wrong* for society to do so. Buchanan (1954a) made an early attack on Arrow's generalization of the concept of individual rationality to collective choice processes focusing precisely on this axiom, and Plott (1972) has extended and generalized this line of criticism. If the transitivity axiom is to earn a place in our constitutional set of constraints on the social choice process, then it must do so by demonstrating that the arbitrary outcomes arising from cyclic preference orderings violate some basic ethical norm. This need not be true. Small committees often resort to random processes such as the flip of a coin, or the drawing of straws to resolve issues of direct conflict. Although obviously arbitrary, the general popularity of random decision procedures to resolve conflictual issues suggests that "fairness" may be an ethical norm that is more basic than the norm captured by the transitivity axiom for decisions of this sort. One might then think of replacing Arrow's notion of collective rationality with the requirement that the social decision process be fair. Transitivity could then be relaxed by simply declaring society indifferent to all choices along the Pareto frontier. Any choice among them will be somewhat arbitrary, but it just might meet with general acceptance. The winners of chess, tennis, and similar elimination tournaments may on occasion be dependent on the particular set of drawings (paths) occurring. This does not seem to detract from the widespread acceptability of this form of tournament for determining the "best" player, however, since the method of determining a sequence of plays is regarded as fair, and the nature of the process precludes the determination of which of the contests were, in fact, path dependent. Thus, it is possible that a social decision process which was intransitive or path dependent, but had additional desirable properties such as fairness, could be widely acceptable. If there is more general agreement concerning these rules than for transitivity or the other consistency properties, the Arrow problem is solved (Kemp, 1954).

[12] Plott (1976, pp. 569–75); Sen (1977a, pp. 71–5).

24.2.2 *Unrestricted domain*

The justification for requiring this axiom is something akin to requiring freedom of choice or expression. Each individual should be free to have any preference ordering he might select and the collective choice process should be capable of reflecting these preferences in accordance with the other axioms. Although freedom of choice strikes a responsive chord, we have seen how quickly conflict can arise when individuals have different preference orderings even over how a given piece of public land is to be used. A set of cyclic preferences is quite possible, and if we also require transitivity, we are well on the way to an impossibility result. It should be obvious that some preference orderings are diametrically opposed to one another. This must follow almost of necessity from Axiom 1, which limits consideration to points along the Pareto frontier, that is, to pure distributional issues. Establishing a committee procedure to resolve these issues, without placing any constraints on the preferences that the individuals can express, seems doomed to failure from the start. Indeed, Saari (1994, p. 327) has observed that the combination of unrestricted domain plus the independence of the irrelevant alternatives axiom allows individual preference orderings to be intransitive. Is it any wonder that the social ordering may violate transitivity?

There are two ways around this problem. One is to replace unrestricted domain with other axioms limiting the types of preference orderings that the collective choice process is capable of reflecting. In the context of public choice, this implies placing constitutional constraints on the types of issues that can come up before the collective. The protection of certain property rights is one example of this type of constraint. Everyone can be a member of the community, but not every preference can be satisfied or necessarily even recorded as part of the collective choice process. The alternative solution is to restrict entry into the community to those having preference orderings that do make collective choices possible.

The first thing to note in this context is that requiring that individual preferences be transitive will *not* solve our problem. Something more, like extremal restriction, is required.[13] Single-peakedness ensures that majority rule produces an outcome, namely, the median, and single-peakedness along with the other four axioms produces a nondictatorial SWF. But this way out of the dilemma requires strict restrictions on both the selection of issues to be decided and the voters to decide them (Slutsky, 1977b). Issues must all be of the one-dimensional variety: the number of guns, the number of schoolbooks. The voters cannot simultaneously consider both the number and kinds of books; and their preferences must be single-peaked in this one dimension. If fate provides voters of this type, these kinds of issues can be resolved by majority rule without violating the other axioms, although we are still left with a plethora of multidimensional issues to resolve in some other way. If some individuals' preferences have multiple peaks, these individuals must somehow be isolated and excluded from the community, or an impossibility result can again emerge.

[13] See Chapter 5.

The single-peakedness and extremal-restriction conditions implicitly introduce a degree of homogeneity of tastes assumption, for there must be a consensus over how the social choices are ordered along some left–right dimension.[14] More generally, the experimental work on majority rule cycles reviewed in Chapter 5 indicates that the probability of a cycle occurring decreases as voter preferences become more "homogeneous," and increases with increasing voter "antagonism" (Plott, 1976, p. 532). These results suggest searching for ways of restricting membership in the polity to those with sufficiently homogeneous or complementary preferences to avoid the impossibility result. The theories of clubs and voting-with-the-feet describe processes by which groups with homogeneous tastes might form. In the absence of externalities across clubs (local communities), and with perfect mobility, free entry, and so on, such a process might avoid the Arrow problem. But, as we have seen, when spillovers exist, some decisions may have to be made by the aggregate population, and the impossibility problem will apply here, even when "solved" in the smaller ones. In such likely circumstances, homogeneity of preferences can be brought about only if individuals adopt or already have a common set of values (Bergson, 1954). Appeals to reason, à la Kant, or uncertainty, à la Rawls (1971) and Harsanyi (1955), are along these lines.

24.2.3 *Independence of irrelevant alternatives*

Of all the axioms, the independence of irrelevant alternatives has been the subject of the most discussion and criticism.[15] In justifying this axiom Arrow (1963, p. 110) made the following argument:

> The Condition of Independence of Irrelevant Alternatives extend the requirements of observability one step farther. Given the set of alternatives available for society to choose among, it could be expected that ideally, one could observe all preferences among the available alternatives, but there would be no way to observe preferences among alternatives not feasible for society . . . clearly, social decision processes which are independent of irrelevant alternatives have a strong practical advantage. After all, every known electoral system satisfies this condition.

Here Arrow defends the axiom in terms of limiting attention to feasible alternatives only, and this objective of the axiom has led Plott (1971, 1976) to restate and rename the axiom specifically in terms of infeasible alternatives. But in his original discussion of the axiom, Arrow presents an example using the rank-order or Borda method discussed in Chapter 7, in which candidates are ranked according to their position in each voter's preferences. In the example Arrow (1963, p. 27) gives, x wins from a slate of x, y, z and w, but draws with z when y is dropped from the list. Thus, under the Borda method the outcome depends on the nature of the full list of candidates. One of Arrow's objectives for invoking the independence axiom would appear to be to eliminate procedures like the Borda method so that "Knowing the

[14] Arrow (1963, p. 80) and Sen (1970a, pp. 166–71).

[15] As noted in n. 3, Arrow's statement of the axiom differs from the one presented here.

social choices made in pairwise comparisons in turn determines the entire social ordering and therefore the social choice function $C(S)$ for all possible environments" (p. 28). Now this is precisely what the independence axiom stated earlier (condition 5) achieves, and it does eliminate procedures like the Borda method from consideration. Thus, our use of this form of the independence axiom would appear to be fully consistent with Arrow's objectives in introducing it.[16] The question then is, what is the normative value to limiting the informational content of collective choice processes in this way?

The outcomes under the Borda procedure and similar schemes depend on the specific (and full) set of issues to be decided. Thus, abandonment of the independence axiom raises the importance of the process that selects the issues to be decided in a way that its acceptance does not. When the choice between x and y can be made by considering voter preferences on only x and y, the rest of the agenda need not be known. This property of the independence axiom has an appealing economy to it, but it is this property that opens the door to endless cycling over these *other* items in the agenda.

By restricting the choice between two alternatives to information on individual rankings of these two alternatives, the independence axiom excludes all information with which one might cardinalize and interpersonally compare utilities (Sen, 1970a, pp. 89–91). It was the desire to establish a welfare function that was not based on interpersonal utility comparisons that first motivated Arrow (1963, pp. 8–11, 109–11). There would appear to be two distinct justifications for wishing to exclude cardinal utility information from a collective choice process. The first is that the measurement of cardinal utilities is difficult and arbitrary, and any process that was based on combining interpersonally comparable, cardinal utilities would be vulnerable to abuse by those making the cardinal utility measurements. This would appear to be Arrow's chief fear (pp. 8–11). It rests on Arrow's view of the collective choice process as one in which information is gathered by public officials who make the actual choices for the collective (pp. 106–8). Allowing these officials to engage in cardinal, interpersonal utility comparisons would vest them with a great deal of discretionary power and might be something to be avoided.

The danger of an abuse of discretionary power does not arise, however, if the cardinal utility information is provided by the voters themselves, as when they take part in the process using, say, the point voting procedure discussed in Chapter 8. Now a different problem arises, however. Such procedures are vulnerable to the strategic misrepresentation of preferences. The independence axiom eliminates not only these strategy-prone procedures, but all voting procedures that are vulnerable to strategizing. This property is sufficiently important to warrant separate treatment.

[16] As Plott (1971, 1976) and Ray (1973) have shown, however, Arrow's original statement of the axiom as given in n. 3 does not exclude the Borda procedure limited to outcomes in the feasible set. It does eliminate the Borda procedure when the ranks are assigned over the set of all possible alternatives, feasible and infeasible, and thus does limit some of this procedure's scope for strategic behavior (Plott, 1976). For additional comment on this axiom, see Bergson (1954), Blau (1972), Hansson (1973), Kemp and Ng (1987), and Saari (1994).

24.3 Strategy-proof social welfare functions

The preceding discussion indicates that an important objective of Arrow in imposing the independence of irrelevant alternatives axiom was to eliminate the possibility of individuals being made better off under a collective decision procedure if they did not state their true preferences as inputs into the collective decision process. Vickrey (1960, pp. 517–19) speculated that immunity to strategic manipulation and satisfying the independence axiom were logically equivalent, and subsequently this insight was rigorously established by Gibbard (1973) and Satterthwaite (1975).

The relationship between independence of irrelevant alternatives (IIA) and strategy proofness (SP) is brought out most clearly by Blin and Satterthwaite (1978).

Strategy proofness (SP): *Let M_i be the message i supplies the voting procedure when she states her true preferences. Let M_i' be any misstatement of i's preferences. Let x be the social outcome from the voting procedure when i states M_i and all other voters j state their true preferences M_j. Let y be the social outcome when i states M_i' and all other voters state their true preferences M_j. Then a voting procedure is strategy proof, if and only if for all possible M_i' there exists no y such that yP_ix.*

Another way to think of strategy proofness is that every profile of true preferences must be a Nash equilibrium under the voting procedure (Blin and Satterthwaite, 1978, p. 257, n. 10).

Blin and Satterthwaite first prove an Arrow-type impossibility theory for the three axioms, nondictatorship (ND), Pareto optimality (PO), and IIA, and two not yet defined axioms, rationality (R) and positive association (PA). R states simply that the voting procedure must define a social preference ordering and subsumes transitivity. PA requires that if x is chosen under one profile of individual preferences, then it must also be chosen under a second profile of preferences that differs from the first only in that x has gone up in one or more individuals' preference orderings.[17]

They then show that the three axioms R, IIA, and PA are equivalent to R and SP. Thus, SP and IIA are not equivalent, but when one demands that the voting process be rational, that is, that it define a consistent social ordering, they come close to being so.

Because R, IIA, and RP are equivalent to R and SP, and it is impossible to have a voting procedure that satisfies R, IIA, RP, ND, and PO, it is impossible to have a voting procedure that satisfies R, SP, ND, and PO. To see the logic of this result, consider a simple example, where we have but two voters (1 and 2) and three alternatives (x, y, z).[18] Each voter can order the three alternatives in six possible ways. Thus, there are 36 possible combinations of the two voters' preference

[17] Note that this axiom is not the same as the positive responsiveness axiom used in May's (1952) theorem on majority rule (discussed earlier in ch. 6). Rather, it resembles nonnegative responsiveness as defined by Sen (1970a, pp. 68–9, 74–7).

[18] With this example we follow the exposition of Feldman (1979, pp. 465–72). Kalai and Muller (1977) show that a strategy-proof SWF exists for a group of $n > 2$, if and only if it exists for a group of two. Thus, a complete proof for a committee of two would suffice for the general case.

Table 24.1. *Possible orderings (6 of 36) of two voters' preferences over three issues*

1	2	1	2	1	2	1	2	1	2	1	2
x	x	x	x	x	y	x	y	x	z	x	z
y	y	y	z	y	x	y	z	y	x	y	y
z	z	z	y	z	z	z	x	z	y	z	x

orderings, of which six are presented in Table 24.1. Voter 1's preferences are the same in all six cases, xP_1yP_1z, and 2's preferences run through the full set of six possible orderings. In the first two combinations or orderings, both voters rank x highest. Thus, by the Pareto principle, x must be the social choice if both voters state either of these sets of preferences. By further application of the Pareto principle, we establish the following restrictions on the social choice for the six combinations of preferences.

$$x \quad x \quad x \text{ or } y \quad x \text{ or } y \quad x \text{ or } z \quad x \text{ or } y \text{ or } z.$$

Voter 1's preferences are the same in all six cases. If 1 honestly states this preference ordering, then any differences in the outcomes that come about must be due to differences in 2's stated preferences. Now consider the third case, where 2's preferences are yP_2xP_2z. This preference ordering in conjunction with 1's must yield either x or y as the social choice to be consistent with the Pareto principle. Suppose from this third case the social outcome were x. Voter 2 prefers y to x under the preferences given in this third case. If they are his true preferences, and the voting rule were such that y would be the outcome if 2 stated any of the preferences 4, 5, or 6, then the procedure would not be strategy-proof; 2 would then state the preference ordering that produced y, given 1's honestly stated ordering. Thus, given that x is the social outcome in case 3, y cannot be the outcome in cases 4, 5, and 6, and we now have the following constraints on the social outcome imposed by strategy proofness.

$$x \quad x \quad x \quad x \quad x \text{ or } z \quad x \text{ or } z.$$

Under the preferences of case 4 (yP_2zP_2x), 2 prefers z to x. Were these 2's true preferences, and z were the social choice for either case 5 or 6, 2 would again have an incentive to misstate his preferences so they appear as in either 5 or 6 when they are really as in 4. Thus, strategy proofness requires x to be the social choice for the pairs of preference orderings in cases 5 and 6. But that implies that x is the social choice when 1's preferences are xP_1yP_1z regardless of what 2's preferences are, which is to say that 1 is a dictator.

Had we assumed that y was the outcome from case 3, we could have shown that nonmanipulation required 2 to be a dictator. The remaining 30 cases can be handled in a similar manner.

The close relationship between strategy proofness and independence of irrelevant alternatives is apparent from this example. In the third combination of individual

preferences depicted in Table 24.1, the two individuals disagree only with respect to whether x is better than y, or the reverse. The independence axiom confines the social choice to using only information from the two individuals' rankings of this pair when choosing the socially preferred outcome. If the social choice process picks x in this situation, it effectively makes 1's preferences dominant over 2's, and 1 becomes the dictator. If y is the social choice, 2 is effectively a dictator.

If the voting procedure's selection of an alternative is sensitive to the voter's full statement of preferences over the 3 or more issues in the issue set, the scope for strategic manipulation of the procedure exists *unless* one voter is treated as a dictator. The dictator has an incentive to be honest, and the preferences of the other voters do not matter. When the voting procedure processes information on only individual, ordinal preferences on issue pairs, as required by the independence axiom, and the procedure is positively responsive, voters will honestly state their true preferences. But information on ordinal preferences on issue pairs does not in general suffice to determine a consistent social preference ordering over the full set of issues. One must make one voter a dictator to ensure transitive social preferences.

The public choice literature builds on the behavioral postulate that individuals rationally and slavishly pursue their self-interests. Whenever the outcome of a voting procedure can be manipulated by cheating, this postulate requires that we assume that voters will cheat – thus, the concern in the public choice literature with finding cheat-proof voting procedures, and the importance of the theorems establishing the impossibility of finding such procedures.

But the negative side to these theorems should not be overdrawn. We saw in Chapter 14 that the rational, self-interest assumption does not give us a very satisfactory predictive theory of voter behavior. Individuals appear to be conditioned to behave in ways that do not fit a narrow definition of self-interested behavior. To what extent individuals who vote out of a sense of "civic duty" would vote strategically is not clear, even if they could figure out what their strategic vote should be.[19]

The more sophisticated voting procedures discussed in Chapter 8 require manipulative strategies that are likely to go beyond the capacities of most voters. The obvious strategy of overloading one's vote points on one's most preferred candidates is curbed in Hylland and Zeckhauser's (1979) version of point voting by the use of a square-root aggregation procedure. The demand revelation process is strategy-proof, although perhaps not Pareto optimal.[20] Voting by veto is strategy-proof, but does not define a social preference ordering.[21] The significance of the impossibility results regarding strategy proofness must be examined in each case. Vernon Smith's (1977) experimental results indicating that students using the auction method of

[19] Cox (1997) presents considerable evidence implying that a small, but nontrivial fraction of citizens do vote strategically in some elections.

[20] The demand revelation process also violates the unrestricted domain assumption by placing certain constraints on individual preferences; for example, they prefer paying less taxes than more. See Sugden (1981, pp. 164–5).

[21] Voting by veto attaches probabilities to the outcomes in the feasible set rather than defining a social ordering over them (Mueller, 1984). In general, probabilistic voting rules that satisfy a positive association condition, as does voting by veto, fare much better with regard to strategy proofness than do deterministic rules. See Gibbard (1977) and Barbera (1977). Note also the similarity between this finding and the results for the political competition models (Chapters 11 and 12).

voting did not behave strategically must again be cited as evidence showing that what one can prove to be a certain hypothetical possibility does not always happen.

24.4 Implications for public choice

The Arrow theorem rests on five axioms that appear to be fairly moderate and reasonable restrictions to place on the collective choice process. The theorem states that no process can exist that satisfies all five axioms simultaneously. In designing a collective choice process, writing our political constitution, we must violate one or more of the axioms – although in so doing we may be able to satisfy the others – and still more to be added to the list.

From a public choice perspective, two promising avenues might be followed out of the Arrow paradox. One is to drop the transitivity axiom and abandon the search for a *best* alternative, *the* social preference. In its place could then be substituted the requirement that the social choice process be fair or accord with some other generally held democratic value. For example, one of the probabilistic voting procedures with desirable normative properties like voting by veto could be substituted for the deterministic ones (see n. 21). Alternatively, if a social ordering must be made, then either the independence axiom or unrestricted domain must be relaxed.

If we continue to interpret these axioms as restrictions on the collective choice process written into the constitution, then these conclusions have the following implications. Axiom 1 limits consideration to points along the Pareto frontier. But a choice from among these involves distributional issues directly, and cycles will occur under any voting process requiring less than full unanimity. Thus, if the majority rule or any other less-than-unanimity rule is chosen, some fair or otherwise generally accepted way for breaking cycles must be included in the constitution.

Relaxing the unrestricted domain assumption to allow only single-peaked preferences does not seem to be a very promising way out of the paradox, since so few issues can realistically be thought of as unidimensional. When collective decisions are restricted to the provision of public goods, the restrictions on preferences that underlie Caplin and Nalebuff's (1988) theorem seem likely to be satisfied, and cycles could be avoided by requiring majorities in excess of 64 percent to pass issues. Some other voting rule would still be needed to deal with redistribution issues.

Alternatively, one can think of designing the constitution in such a way as to allow for the revelation of preferences for public goods via voluntary association in private and local clubs. This solution solves the problem by imposing a form of unanimity condition, but again leaves aside all distributional considerations, and the problems of resolving differences of opinion on global public goods.

Where strategic behavior is not a problem, one of the procedures that gathers information on the voters' preferences over the full set of alternatives, like the Borda procedure or point voting, can be used. As we noted in Chapter 8, however, the normative properties of these procedures depend heavily on what issues are allowed into the decision set. Thus, relaxing either the unrestricted domain assumption or independence of irrelevant alternatives raises questions as to what issues are to be decided, who is to decide, and of those who decide, which preferences shall be

weighed and with what weights. Such choices directly or indirectly involve interpersonal utility comparisons and must rest on some additional value postulates which, if explicitly introduced, would imply specific interpersonal utility comparisons. The latter cannot be avoided if a preferred social choice is to be proclaimed.[22]

We close our discussion of the Arrow axiomatic SWF at the same point we were at with the Bergson-Samuelson real-valued SWF.

Bibliographical notes

The difference between Arrow's SWF and the Bergson-Samuelson SWF has been the subject of much discussion (Arrow, 1963, pp. 23–4; Samuelson, 1967; Sen, 1970a, pp. 33–6).

Numerous books and articles survey and extend the impossibility result first established by Arrow. See, in particular, Riker (1961, 1982b), Rothenberg (1961), Arrow (1963, ch. 8), Sen (1970a, 1977a,b, 1999), Pattanaik (1971, 1997), Taylor (1971), Fishburn (1973), Plott (1976), Kelly (1978), MacKay (1980), Suzumura (1983), and Saari (1994).

[22] Kemp and Asimakopulos (1952), Hildreth (1953), Bergson (1954), Sen (1970a, pp. 123–5, 1974, 1977b).