10 ARROW'S IMPOSSIBILITY THEOREM

INTRODUCTION

We have now looked at several ways to answer the question “When is one alternative socially preferred to another?” Each of the answers has been somewhat disappointing. The Pareto criterion is incomplete; the Kaldor criterion is possibly inconsistent. The Scitovsky criterion reduces to a question of optimality vs. non-optimality. The Samuelson criterion might be completely devoid of content. The fairness and Rawls criteria might be inconsistent with the Pareto criterion. Majority voting might generate cycles, and the single-peakedness condition, which forces transitivity, is quite restrictive.

Our goal throughout has been to discover an unerring rule for generating rational social preferences, rational in the sense that an individual’s preferences are rational. That is, we have been looking for a rule that generates complete and transitive social preferences, or, at least, complete and acyclic social preferences. But our series of disappointments raises some questions: Does a foolproof method exist for constructing complete and transitive social preference relations? Does a foolproof method exist for constructing complete and acyclic social preference relations? Does a foolproof method exist for finding best social alternatives? In this chapter we construct a simple model to answer the first question, and we briefly discuss the answer to the second question.

Does a foolproof method exist for constructing complete and transitive social preference relations? The answer to this question clearly depends on what we mean by foolproof. We will impose formal requirements on the method for constructing social preferences, requirements that allow a definite answer. The list of requirements, and the answer to the question, were developed by Kenneth Arrow around 1950, and the answer, to which we shall soon turn, is called Arrow’s Impossibility Theorem.

THE MODEL

We now assume for the sake of simplicity that there are only two individuals, and three social alternatives x, y, and z. Also, we suppose that no individual is ever indifferent between any two alternatives. As usual, xPy means i prefers x to y. Individual i’s preference ordering is assumed to be complete and transitive. Since there are only three alternatives and indifference is disallowed, there are only six ways individual 1 can order the alternatives. He can prefer x to y to z, or he can prefer x to z to y, or he can prefer y to x to z, and so on. The same is true of individual 2. Therefore, if any preference ordering for 1 or 2 is allowed, there are exactly $6 \times 6 = 36$ different constellations of individual preferences, or preference profiles, possible in this small society. Table 10–1 includes them all.

Each cell in this table shows a possible pair of rankings of the three alternatives by individuals 1 and 2. On the left side the alternatives are ordered, from top to bottom, according to person 1’s preferences, and on the right side they are ordered according to 2’s preferences. For example, the 1st row, 2nd column cell has 1 preferring x to y to z, and 2 preferring x to z to y.

Our concern here is whether or not there is a foolproof rule to transform any cell in Table 10–1 into a social preference relation. Such a rule is called a collective choice rule. A collective choice rule takes preference profiles and produces social preferences.

Let R stand for a social preference relation, so xRy means x is socially at least as good as y. P is the corresponding strict social preference relation: xPy means x is socially preferred to y; i.e., xRy and not yRx. (Don’t confuse this R with the Rawls criterion, or this P with the Pareto criterion.) Finally, I is the social indifference relation: xIy means x and y are socially indifferent; i.e., x Ry and yRx.
In the next section we will list five plausible requirements that will be imposed on the collective choice rule. Taken together, these five requirements define what we mean by foolproof.

**REQUIREMENTS ON THE COLLECTIVE CHOICE RULE**

1. **Completeness and transitivity.** The social preference relations generated by a collective choice rule must be complete and transitive. If some preference profile is transformed into a particular \( R \), then for any pair of alternatives \( x \) and \( y \), either \( xRy \) or \( yRx \) must hold, and for any triple \( x, y, z \) if \( xRy \) and \( yRz \) must imply \( xRz \). The requirement says that a collective choice rule must always permit comparisons between two alternatives, and that social preferences must have the nice transitivity property assumed for an individual’s preferences.

We have examined collective choice rules that don’t generate complete and transitive social preference relations. The Pareto criterion gives incomplete social rankings: if \( 1 \) prefers \( x \) to \( y \) and \( 2 \) prefers \( y \) to \( x \), the two alternatives are Pareto noncomparable. Majority voting gives nontransitive social rankings. As an example in a two-person case (where the voting cycle of Condorcet cannot be generated), suppose \( 1 \) prefers \( x \) to \( z \) to \( y \), and \( 2 \) prefers \( y \) to \( x \) to \( z \). If a vote is taken between \( x \) and \( y \), there is a tie (1 votes for \( x \), and 2 votes for \( y \)). If a vote is taken between \( y \) and \( z \), there is again a tie. According to majority voting, \( x \) and \( y \) are socially indifferent, and \( y \) and \( z \) are socially indifferent. Transitivity would require that \( x \) and \( z \) be socially indifferent. But in a vote between \( x \) and \( z \), \( x \) gets 2 votes and \( z \) none; so \( x \) beats \( z \), and transitivity fails.

2. **Universality.** A collective choice rule should work no matter what individual preferences happen to be. This means that the rule should give us a social preference ordering for every cell in Table 10–1, not just for the easy ones, like the ones where there is unanimous agreement.

Universality is a significant requirement. It precludes the assumption of single-peakedness, since it says that the collective choice rule must work for all preference profiles, not just the ones where utility functions have single peaks. Why should it be imposed?

First, it is difficult to see where to draw the line between permissible and impermissible individual preferences. Which cells in Table 10–1 should be disallowed or ignored? How much diversity can be expected in society? Is there so much conflict that the very idea of social welfare becomes implausible? There are no easy answers to these questions. Second, the theorem we will prove remains valid even when the universality requirement is substantially weakened, and we will indicate how much it can be weakened in a subsequent section.

3. **Pareto consistency.** A collective choice rule should be consistent with the Pareto criterion. For any pair of alternatives \( x \) and \( y \), if both individuals prefer \( x \) to \( y \), \( x \) must be socially preferred to \( y \).

Pareto consistency is a very mild requirement for a collective choice rule. One would not expect it to hold in societies which are ruled by external forces; in which, for example, everyone prefers lust and gambling, on the other hand, to chastity and frugality on the other; but where, according to a
Holy Book, the social state of chastity and frugality is preferable to the social state of lust and gambling. Economists naturally would recommend lust and gambling.

On a more serious note, let’s recall that the fairness and Rawls criteria could produce results contrary to the Pareto criterion. In our view, this fact is an indictment of fairness and Rawls, not of Pareto. We take Pareto consistency to be fundamental.

4. Non-dictatorship. A collective choice rule must make no one a dictator. Individual i is said to be a dictator if his wishes prevail, no matter how j feels; that is, if xP_i y implies xP_j y for all x and y, irrespective of P_j. Ruling out dictatorship does not mean that it is never possible to have xP_i y implying xP_j y for all x and y. Obviously, if both people agree on the rankings of all alternatives (so that P_1 = P_2, as in the diagonal cells of Table 10-1), then it is perfectly reasonable to have the social preference relation agreeing with 1’s (and 2’s) preference relation, and in fact, the Pareto consistency requirement makes such agreement necessary. Nondictatorship simply says that 1 (or 2) must not always prevail, no matter how 2 (or 1) happens to feel.

5. Independence of irrelevant alternatives. If people’s feelings change about some set of irrelevant alternatives, but do not change about the pair of alternatives x and y, then a collective choice rule must preserve the social ordering of x and y. The social preference between x and y must be independent of individual orderings on other pairs of alternatives. (We should note that this formulation of independence differs slightly from Arrow’s original formulation.)

Independence is the subtitle of the five requirements, and it takes some explanation. Suppose society prefers x to y when z is a third alternative lurking in the wings. Next suppose everyone suddenly changes his mind about the desirability of z, but no one changes his mind about x vs. y. The independence requirement says that, if society is deciding on the relative merits of x and y, and only those two, it must still prefer x to y.

The standard example of an otherwise-nice collective choice rule that violates independence is weighted voting. This type of rule was first analyzed in 1781 by Jean-Charles de Borda, in his Mémoire sur les Élections au Scrutin, and it is consequently also called de Borda voting. It works as follows. Each person reports his preference relation, his rank ordering. A first place in a rank ordering is assigned a certain fixed weight, a second place is assigned a (smaller) fixed weight, a third place is assigned a (yet smaller) fixed weight, and so on. (In the two-person, three-alternative model of this chapter we have no ties, no cases of indifference, to worry about.) The weights that each alternative gets from each person are summed, and the social preference relation is derived from the sums of the weights.

For instance, suppose person 1 prefers z to x to y, while person 2 prefers y to x to z. Suppose a person’s first choice gets a weight of five points, a second choice gets four points, and a third choice gets one point. (The weights are obviously arbitrary. It is common to use equally spaced weights, like 3, 2, and 1, but there is no logically compelling reason to do so. You may construct an example similar to this one using the common weighting scheme.) Now alternative x gets 4 + 4 = 8 points, alternative y gets 1 + 5 = 6 points, and alternative z gets 5 + 1 = 6 points. Therefore, for this preference profile, x is socially preferred to y according to the weighted voting rule.

However, suppose person 1 becomes disillusioned with alternative z, and his preference ordering changes to x over y over z. If the voting is repeated, x gets 5 + 4 = 9 points, y gets 4 + 5 = 9 points, and z gets 1 + 1 = 2 points. Therefore, given this new preference profile, x is socially indifferent to y. Society has become indifferent between x and y, even though neither person has changed his feelings about x and y! Consequently, weighted voting violates the independence requirement.

APPLYING THE REQUIREMENTS

At this stage we shall apply requirements 1, 2, 3 and 5 to Table 10-1. The applications should clarify the meanings of the four requirements. They will also lay the groundwork for the proof of Arrow’s Impossibility Theorem.

The completeness and transitivity and the universality requirements say that, when applied to any cell of Table 10-1, a collective choice rule must generate a complete and transitive social ordering.

The Pareto consistency requirement says a collective choice rule must respect unanimous opinion: If both 1 and 2 prefer one alternative to another, then society must also prefer the one to the other. For example, given the preference profile of the first row, second column cell of Table 10-1, the Pareto requirement says x must be socially preferred to y and x must be socially preferred to z. We must have xP_y and xP_z. Application of Pareto consistency over the entirety of Table 10-1 gives rise to Table 10-2.

Each cell of this table is produced by applying Pareto consistency to the corresponding cell of Table 10-1, and therefore, any rule for generating social preferences must be entirely consistent with Table 10-2.
### Table 10-2

<table>
<thead>
<tr>
<th>xPy</th>
<th>xPy</th>
<th>xPz</th>
<th>yPz</th>
<th>xPz</th>
</tr>
</thead>
<tbody>
<tr>
<td>yPz</td>
<td>xPz</td>
<td>xPz</td>
<td>yPz</td>
<td>xPz</td>
</tr>
</tbody>
</table>

### Arrow's Impossibility Theorem

Let's be specific about those areas of agreement. Independence requires that all the cells in Table 10-1 where \( xP_1y \) and \( yP_2x \) must yield identical social rankings of \( x \) and \( y \). Similarly, all the cells where \( yP_1x \) and \( xP_2y \) must yield identical social rankings of \( x \) and \( y \). There is no presumption, however, that the social ranking of \( x \) and \( y \) on the \( xP_1y \) and \( yP_2x \) cells need be the same as the social ranking of \( x \) and \( y \) on the \( yP_1x \) and \( xP_2y \) cells. Such a neutrality condition is unnecessary for the proof of the Impossibility Theorem, although it is intuitively appealing and useful in other contexts.

Independence also implies these areas of agreement. All the cells of Table 10-1 where \( xP_1z \) and \( zP_2x \) must give rise to identical social rankings of \( x \) and \( z \); all the cells where \( zP_1x \) and \( xP_2z \) must give rise to identical social rankings of \( x \) and \( z \); all the cells where \( yP_1z \) and \( zP_2y \) must give rise to identical social rankings of \( y \) and \( z \); and, finally, all the cells where \( zP_1y \) and \( yP_2z \) must give rise to identical social rankings of \( y \) and \( z \).

All of this information can be incorporated in a third table. Table 10-3a indicates where the social rankings of \( x \) and \( y \) must agree because \( xP_1y \) and \( yP_2x \) in all the \( X \)'d cells, and where the social rankings of \( x \) and \( y \) must agree because \( yP_1x \) and \( xP_2y \) in all the \( O \)'d cells. Tables 10-3b and 10-3c show the areas of agreement which arise from applications of the independence requirement to the social preferences between \( x \) and \( z \), and \( y \) and \( z \), respectively.

### Table 10-3a

<table>
<thead>
<tr>
<th></th>
<th>x</th>
<th>x</th>
<th>x</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>x</td>
<td>x</td>
<td></td>
</tr>
</tbody>
</table>

The crossed cells all produce the same \( x \cdot y \) social rankings. The circled cells all produce the same \( x \cdot y \) social rankings (which need not be the same as in the crossed cells).
TABLE 10-3b & c

<table>
<thead>
<tr>
<th></th>
<th>X</th>
<th>X</th>
<th>X</th>
</tr>
</thead>
<tbody>
<tr>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
</tbody>
</table>

The crossed cells all produce the same x-z social rankings. The circled cells all produce the same x-z social rankings (which need not be the same as in the crossed cells).

(b)

<table>
<thead>
<tr>
<th></th>
<th>X</th>
<th>X</th>
<th>X</th>
</tr>
</thead>
<tbody>
<tr>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
</tbody>
</table>

The crossed cells all produce the same y-z social rankings. The circled cells all produce the same y-z social ranking (which need not be the same as in the crossed cells).

(c)

With these preliminaries out of the way, we can turn to a truly remarkable theorem.

ARROW'S IMPOSSIBILITY THEOREM

At least since the time of Condorcet and de Borda in the eighteenth century, people have been concerned with the properties of rules for making social choices, election rules in practice, collective choice rules in theory. Does there exist a foolproof rule for discovering, or for defining, social prefer-

ences? Arrow showed that, if foolproof means consistent with the five requirements above, the answer is No.

We now turn to a formal statement and proof of the theorem.

**Arrow's Impossibility Theorem.** Any collective choice which is consistent with the requirements of (1) completeness and transitivity, (2) universality, (3) Pareto consistency, and (4) independence of irrelevant alternatives, makes one person a dictator. Therefore, there is no rule which satisfies all five requirements.

**Proof.** We start by looking at the preference profile of the first row, second column cell of Table 10-1. For these preferences Pareto consistency requires \(xPy\) and \(xPz\) (Table 10-2). There are three and only three complete and transitive social preference orderings which satisfy \(xPy\) and \(xPz\). They are:

1. \(xPy, xPz\) and \(yPz\)
2. \(xPy, xPz\) and \(zPy\)
3. \(xPy, xPz\) and \(yPz\)

Each of these three possibilities will be considered in turn.

**Case 1:** \(yPz\). First a word about strategy. The Pareto consistency requirement tells a lot about what social preferences must be, but it leaves a lot unsaid. Table 10-2 is full of blank and partially blank spaces. We will now show how all the blanks can be filled in by repeatedly applying the independence and transitivity requirements.

If \(yPz\) holds in the first row, second column cell, then independence (Table 10-3c) requires that \(y\) be socially preferred to \(z\) whenever individual preferences about \(y\) and \(z\) are the same as they are in that cell. Therefore, \(yPz\) holds in all the cells indicated in Table 10-4a. (The cells that provide crucial steps in the proof are numbered 1 and 2.)

Now consider the first row, fifth column cell, or cell number 2 in Table 10-4a. Pareto consistency (Table 10-2) requires that \(xPy\) here, but \(xPy\) and \(yPz\) implies \(xPz\), by transitivity. So in this cell we must also have \(xPz\).

But if \(xPz\) holds in cell number 2, then independence (Table 10-3b) requires that \(x\) be socially preferred to \(z\) whenever individual preferences about \(x\) and \(z\) are the same as they are in that cell. Therefore, \(xPz\) holds in all the cells indicated in Table 10-4b. (The cells that provide crucial steps in the proof are numbered 2 and 3.)

Now we have \(xPz\) in cell 3. We again invoke Pareto consistency and transitivity to conclude that \(xPy\) must hold in cell 3 as well. But this
allows us to use independence again, (Table 10-3a), to fill in eight more bits of information.

The filling-in process is repeated four more times. You can complete this part of the argument using the sequence of crucial cells indicated in Table 10-4c.

This filling-in process produces six diagrams like Tables 10-4a and 10-4b. Each one shows nine identical social preferences. If the information contained in all these diagrams is added to the information of

Table 10-2, the result is the pattern of social preferences of Table 10-5.

But the social preferences shown in Table 10-5 are identical to person 1's preferences. Therefore, in Case 1, 1 is a dictator. He gets his way, no matter how 2 feels.

Case 2: zPy. If zPy holds in the 1st row, 2nd column cell, an argument analogous to the one above establishes that 2 is a dictator. The argument is left as an exercise.

Case 3: yfz. Suppose yfz holds in the first row, second column cell. Then by independence of irrelevant alternatives (Table 10-3c) yfz must
also hold in the third row, second column cell, as well as the fourth row, fifth column cell. By Pareto consistency (Table 10–2), z must be socially preferred to x in the latter cell. Now by transitivity, \( y \succ z \) and \( z \succ p \) implies \( y \succ p \), for the fourth row, fifth column cell. By independence again (Table 10–3a), \( y \succ p \) in the fourth row, fifth column cell implies \( y \succ p \) in the third row, second column cell. Using transitivity again, \( y \succ z \) and \( y \succ p \) implies \( z \succ p \) in this cell. However, this contradicts Pareto consistency (Table 10–2), which says that \( x \succ p \) holds here. Therefore, Case 3 is impossible.

The proof of the theorem is now complete, for it has been shown that requirements 1, 2, 3, and 5 together imply that either

i. 1 is a dictator, or

ii. 2 is a dictator.

Q.E.D.

RELAXING THE UNIVERSALITY REQUIREMENT

We said near the beginning of this chapter that the universality requirement, which demands that a collective choice rule work for any preference profile, is overly strong. This section will indicate why. In fact, the construction of a sensible collective choice rule remains impossible even if a large number of possible preference profiles are excluded.

To show that lots of preference profiles might be dispensed with, without affecting the theorem, we will just count the cells that are crucial in our proof. In Case 1, we use the cells numbered 1–6 in Table 10–4c. In Case 2, we need another six crucial cells, but the first crucial cell is again the first row, second column cell, which we have already counted. Thus far, we have \( 6 + 5 = 11 \) cells. In Case 3, we use three crucial cells, but one is the familiar first row, second column cell, which has been counted. Therefore, the total number of crucial cells for the purposes of our proof is \( 11 + 2 = 13 \). And the Arrow Impossibility Theorem holds even if any or all of the remaining cells are discarded. Incidentally, the crucial cells we have used are not the only ones which can establish the theorem—other sequences of steps can be used to prove it. But they are a full set, in the sense that they will do the job, and so long as they are all retained, any or all of the other cells are disposable.

In short, with a properly chosen set of thirteen preference profiles, out of the total set of thirty-six, Arrow's Theorem can be established. So universality is really a much stronger assumption than is needed to prove the theorem.

ARROW'S IMPOSSIBILITY THEOREM

Since Arrow published his seminal paper in 1950, a vast literature has grown on the Impossibility Theorem and related topics. It's interesting to think about why. The theorem provides an unambiguous answer to the question "Is there a foolproof way to derive complete and transitive social preference relations?" The answer is No. This clearly negative result casts doubts on all assertions that there is a "general will," a "social contract," a "social good," a "will of the people," a "people's government," a "people's voice," a "social benefit," and so on and so forth. That is, it casts doubts on all notions that explicitly or implicitly attribute preferences to society that are comparable to preferences for an individual. Therefore, it casts doubts on vast areas of twentieth century social thought. Of course any theorem that casts so much doubt will generate a lot of responses.

There are many possible reactions to Arrow's Theorem. The first, and perhaps most obvious reaction is this. It is quite silly in the first place to think that there might be social preferences that are analogous to individual preferences. It is nonsense to talk about social preferences since society itself is nothing more than a collection of individuals, each with his own interests. The idea that a motley collection of individuals should have social preferences that are like an individual's preferences is just an example of illegitimate reasoning by analogy. To attribute the characteristics of an individual to a society commits the logical error of personification. Arguments like these have been made by James Buchanan, Charles Plott, and others.

Now this line of reasoning could be pursued further. For instance, the idea that anyone should be interested in Pareto optimality is also silly, since each person \( i \) just wants to maximize his own utility \( u_i \). Person \( i \) couldn't care less about optimality. Further, the idea that a government official might be interested in pursuing the public good is equally silly, first because the public good is an empty idea, and second because even if there were a public good to be pursued, public official \( i \) just wants to maximize his utility function \( u_i \), and what's good for the public will often differ from what's good for the public official. (See Anthony Downs' An Economic Theory of Democracy.) Further, the idea that an economist or a political scientist might be interested in instructing the citizenry or electorate, or reforming public officials, is also silly, first because there is no public good, second because the citizens and/or the officials are only trying to maximize their own utility functions, and third because the economist or the political scientist is also simply interested in maximizing his \( u_i \). And why should anyone listen to such an obviously self-serving advisor or reformer?
So this first reaction to Arrow’s Theorem is logically attractive, but it can lead to varieties of nihilism that are unappealing to some people.

The second type of reaction to Arrow’s Theorem accepts the legitimacy of the basic idea of social preferences, and attacks one or more of Arrow’s five requirements. Let’s very briefly touch on some of these lines of attack.

First, the completeness requirement might be jettisoned. For instance, some people are quite satisfied with the Pareto criterion alone, and hold that if \( x \) and \( y \) are Pareto noncomparable, only an act of God can decide between them. There is no way that a reasonable person, or government, can decide, and reasonable people or governments have no business trying to do so. One implication of this line of reasoning is that a Pareto optimal status quo, brought about, perhaps, by a competitive market mechanism, should be left untouched—government officials should not waste their time (and tax dollars) scheming about how to redistribute wealth. Nor should economists and political scientists.

Second, the transitivity requirement might be dropped, or weakened to quasi-transitivity, or acyclicity. If transitivity is dropped entirely, majority voting becomes acceptable, and its advocates simply hope embarrassing voting cycles do not arise. If cycles are potentially there, they might be suppressed by clever committee chairmen, by agenda rules, or by some other deus ex machina.

Or, transitivity might be weakened to quasi-transitivity. If this is done, however, a new version of Arrow’s Impossibility Theorem rears its head. In this version, due to Allan Gibbard, the requirements of completeness and quasi-transitivity, universality, Pareto consistency, and independence together imply that there must exist an oligarchy, rather than a dictatorship. That is, there must exist a group of people \( G \) such that (1) if \( xPy \) for all \( i \) in \( G \), then \( xPy \), and (2) if \( xPy \) for some \( i \) in \( G \), then \( xPy \). Acting together, members of the oligarchy can force a social preference for \( x \) over \( y \) and, acting apart, each member of the oligarchy has veto power over a \( y \) which he regards as inferior to \( x \). If \( G \) has only one member, then the oligarchy is a dictatorship. If \( G \) has two members, it is a dumvirate; if it has three members, it is a triumvirate, and so on. \( G \) might include everyone. Consider, for instance, A. K. Sen’s Pareto-extension rule, which is defined as follows. Let \( xPy \) if \( y \) is not Pareto superior to \( x \). This rule satisfies completeness and quasi-transitivity, universality, Pareto consistency, and independence. It makes the whole group an oligarchy. And, we should note, it doesn’t provide much information. It says, for example, that all Pareto optimal allocations are equally good.

If transitivity is weakened to acyclicity, there arises yet another version of Arrow’s Impossibility Theorem, a version due to Donald Brown. This result says the requirements of completeness and acyclicity, universality, Pareto consistency, and independence together imply that there must exist what Brown calls a collegial polity, rather than an oligarchy or a dictatorship. In a collegial polity, there may be several groups \( G \) who can force a social preference for \( x \) over \( y \), that is, several \( G \)'s such that if \( xPy \) for all \( i \) in \( G \), then \( xPy \). These groups have a nonempty intersection, and this intersection, this set of people who belong to all the powerful \( G \)-groups is called a collegium, which we abbreviate \( C \). Now \( C \) need not be able to force the social preference of \( x \) over \( y \) by itself, but any group that can force that preference must include \( C \). \( C \) is thus the elite of the powerful. As an example, suppose there are five people, and the collective choice rule defines \( xPy \) if persons 1 and 2 plus any one other person prefer \( x \) to \( y \). Otherwise, \( yRx \). Then the powerful \( G \)-groups are \( \{1,2,3\} \), \( \{1,2,4\} \) and \( \{1,2,5\} \). That is, any of these groups can force a social preference for \( x \) over \( y \). The elite among the powerful, the collegium, is \( \{1,2\} \). However, the collegium \( \{1,2\} \), acting alone, cannot force a social preference for \( x \) over \( y \).

A third reaction to Arrow’s Theorem is to drop the universality requirement. This is the approach of the single-peakedness mode of analysis, and similar analyses of restrictions of allowable preference profiles. We have already seen in Chapter 9 and this chapter that this tack is disappointing.

A fourth, and for our purposes final, way is to drop the independence requirement. There are at least two reasons why one might want to move in this direction. The first is that the independence requirement is, of all the Arrow requirements, the least intuitive, the least compelling, the least understandable, and these weaknesses suggest it should be sacrificed. The second is that the independence requirement depends on there being more than one preference profile, it depends on changes in individuals’ preferences. Why not constrain the entire discussion to a fixed preference profile? Why not say: “Look, we have so many people with particular preferences that are given. How might we aggregate these given preferences? What might or might not happen if preferences change is of no particular interest, because we want to aggregate the fixed preferences of our given population.”

What is wrong with this position? The essential problem here is that with the independence requirement we are forced to admit only one type of collective choice rule—dictatorship, but without it we have an unlimited set of admissible collective choice rules, a real embarrassment of riches. For instance, suppose we forget about independence, suppose for simplicity that no one is ever indifferent between two alternatives, and suppose the total number of alternatives is \( k \). Let \( \{a_1, a_2, a_3, \ldots, a_k\} \) be any \( k \) numbers satisfying \( a_1 > a_2 > a_3 > \ldots > a_k \). Consider the generalized weighted voting rule, or generalized de Borda rule, in which the weight assigned a first choice is \( a_1 \), the weight assigned a second choice is \( a_2 \), and so on. Then this collective choice rule is perfectly well behaved; it satisfies completeness.
and transitivity, universality (which might be moot if we are really concerned with a fixed preference profile) and Pareto consistency. Independence has been jettisoned. And the rule is most certainly nondictatorial, nonoligarchic and noncollegial. But the problem here is that the outcome, the social preference relation, depends on the actual magnitudes of the weights \((a_1, a_2, \ldots, a_k)\), and that a different set of weights will generally give a different social preference relation. And there are infinitely many ways to choose the weights! So the resulting social preference relation is arbitrary, insofar as the particular weights are arbitrary.

How is a (social) choice of a set of weights going to be made? How can we decide if one set of weights \((a_1, a_2, a_3, \ldots, a_k)\) is socially preferable to another set of weights \((b_1, b_2, b_3, \ldots, b_k)\)? The problem of social preferences has not been solved in this case, it has only been thrush back onto the choice of the weights.

**EXERCISES**

1. Suppose there are three people and four alternatives. Assume 1 prefers \(w\) to \(x\) to \(y\) to \(z\), 2 prefers \(y\) to \(z\) to \(w\) to \(x\), and 3 prefers \(z\) to \(x\) to \(y\) to \(w\). Find weights \((a_1, a_2, a_3, a_4)\) so that weighted voting indicates the social preference is \(y\) over \(z\) over \(x\) over \(w\). That is, \(y\)'s total weighted vote is the highest, \(z\)'s is the second highest, and so on. Next find weights \((b_1, b_2, b_3, b_4)\) so that weighted voting indicates the social preference is \(z\) over \(x\) over \(y\) over \(w\).

2. Suppose there are four people, and no one is ever indifferent between two alternatives. One of the alternatives is special, and is labeled \(x_0\). Consider a collective choice rule that works as follows. If person 1 prefers \(x\) to \(y\), and \(y\) isn't the special alternative, then \(x\) is socially preferred to \(y\). That is, \(xP_1y\) implies \(xPy\), if \(y \neq x_0\). (So person 1 is a dictator except when the special alternative is at stake.) If persons 1 and 2, plus at least one other person, prefer \(x\) to \(x_0\), then \(x\) is socially preferred to \(x_0\). (So person 1 needs person 2 plus someone else to overrule the special alternative.) In all other cases, \(yRx\).
   a. Show that this rule satisfies Pareto consistency and independence.
   b. Show with specific examples of preferences that the rule is not quasi-transitive.
   c. Show that it must be acyclic.
   d. Identify the collegium.
   e. If all members of the collegium prefer \(x\) to \(y\), does it necessarily follow that \(xPy\)?

3. Prove Case 2 of Arrow's Impossibility Theorem.

**SELECTED REFERENCES**

(Articles marked with an asterisk (*) are mathematically difficult.)


   Arrow's 1950 paper is the source of the enormous literature on the Impossibility Theorem. (By now there are many hundreds of published papers on the topic.) After briefly surveying dictatorship, rule by convention, majority rule, and compensation criteria, Arrow turns to the question: "Can we find other methods of aggregating individual tastes which imply rational behavior on the part of the community and which will be satisfactory in other ways?" The answer is No.


   Arrow's excellent monograph extends and supplements his original 1950 article. The monograph is as valuable for its treatment of related topics as it is for its proof of the Impossibility Theorem. For instance, Arrow discusses compensation criteria at great length; he has a chapter on preference similarity that deals with unanimity, single-peakedness, and in the last chapter he discusses some of the issues raised in the literature on social choice since the publication of his 1950 article (and the first, 1951 edition of this monograph). The last section of the last chapter is especially interesting, since it is about whether or not rationality is a property that ought to be attributed to society.


   See the references section at the end of Chapter 1.


   Plott provides an interesting and relatively nontechnical survey of Arrow's Impossibility Theorem topics in the first part of this paper. In the second, he surveys axiomatic characterizations of collective choice rules. These are statements of the form "if a rule has properties A, B and C, then it must be such-and-such."


   Chapter 3 is a literary treatment of Arrow's results, 3* is a mathematical treatment and a general proof. Chapters 4 and 4* are about weakening the transitivity requirement.


   This article surveys the technical literature on collective choice rules that generate binary social preference relations, and collective choice rules that generate choice functions. (A choice function indicates what alternatives are best from among any subset of alternatives.) Impossibility results are given for both types of rules.