CHAPTER 6

SOCIAL CHOICE AND WELFARE

With only few exceptions, we have so far tended to concentrate on questions of 'positive economics'. We have primarily been content to make assumptions about agents' motivations and circumstances, and deduce from these the consequences of their individual and collective actions. In essence, we have characterised and predicted behaviour, rather than judged it or prescribed it in any way. In most of this chapter, we change our perspective from positive to normative, and take a look at some important issues in welfare economics. At the end of the chapter we return to positive economics and consider how individuals motivated by self-interest make the problem of social choice doubly difficult.

6.1 THE NATURE OF THE PROBLEM

When we judge some situation, such as a market equilibrium, as 'good' or 'bad', or 'better' or 'worse' than another, we necessarily make at least implicit appeal to some underlying ethical standard. People often differ in their systems of ethics and so differ in their judgements on the merits of a given situation. This obvious fact need not discourage us nor make us despair that normative economics is all 'just a matter of opinion'. On the contrary, there is such a thing as consistency in reasoning from premises to conclusions and so to prescriptions. Welfare economics helps to inform the debate on social issues by forcing us to confront the ethical premises underlying our arguments as well as helping us to see their logical implications.

Viewed broadly, our goal in much of this chapter is to study means of obtaining a consistent ranking of different social situations, or 'social states', starting from well-defined and explicit ethical premises. On the level of generality at which we shall work, a 'social state' can be just about anything: the election of a particular candidate to a political office, a particular way of dividing a pie among a group of people, adoption of a market-oriented form of organising society, or a particular way of distributing society's resources among its members. A social choice problem arises whenever any group of individuals must make a collective choice from among a set of alternatives before them.

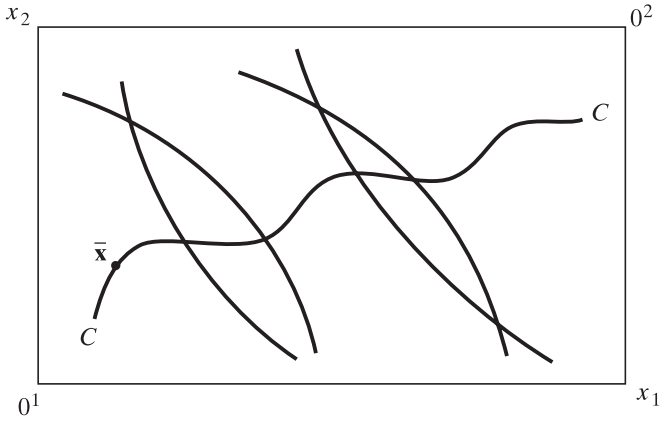


Figure 6.1. The distribution problem.

To make things a bit more concrete for just a moment, let us consider the distribution problem in a simple, two-good, two-person Edgeworth box economy, like the one depicted in Fig. 6.1. There, each point in the box represents some way of dividing society's fixed endowment of goods between its two members, so we can view each point in the box as one of the (mutually exclusive) alternate social states we could achieve. Each agent has his or her own preferences over these alternatives, and clearly these preferences are often at odds with one another. The social choice problem involved is easy to state. Which of the possible alternative distributions is best for society?

Although easy to state, the question is hard to answer. Perhaps without too much disagreement, points off the contract curve can be ruled out. Were one of these to be recommended as the best, it would be easy to find some other point on the contract curve that everyone prefers. Because it would be hard to argue with such unanimity of opinion, it is probably safe to say that our search for the best alternative ought to be restricted to the Pareto-efficient ones.

But which of these is best? Many will find it easy to say that wildly unequal alternatives such as $\bar{\mathbf{x}}$ must also be ruled out, even though they are Pareto efficient. Yet in doing so, appeal is being made to some additional ethical standard beyond the simple Pareto principle because that principle is silent on the essential question involved: namely, how may we trade off person 2's well-being for that of person 1 in the interests of society as a whole? In trying to make such trade-offs, does *intensity* of preference matter? If we think it does, other questions enter the picture. Can intensity of preference be known? Can people tell us how strongly they feel about different alternatives? Can different people's intense desires be compared so that a balancing of gains and losses can be achieved?

The questions are many and the problems are deep. To get very far at all, we will need to have a systematic framework for thinking about them. Arrow (1951) has offered such a framework, and we begin with a look at his path-breaking analysis of some of these problems.

6.2 SOCIAL CHOICE AND ARROW'S THEOREM

The formal structure we adopt is very simple and very general. There is some non-empty set X of mutually exclusive social states under consideration. While just about everything we do in this chapter can be accomplished whether the set X is finite or infinite, to keep things simple we will sometimes assume that X is finite and other times assume that it is infinite. We will be sure to let you know which of these we are assuming at all times. Society is composed of N individuals, where $N \geq 2$. Each individual i has his own preference relation, R^i, defined over the set of social states, X, with associated relations of strict preference, P^i, and indifference, I^i. Being a preference relation, each R^i is *complete* and *transitive*. Intuitively, we require nothing but that people be able to make binary comparisons between any two elements in X, and that those comparisons be consistent in the sense of being transitive. The set X has been defined very broadly, so keep in mind that its elements may range from the purely mundane to the purely spiritual. The relations R^i, therefore, also must be broadly construed. They need not merely reflect selfish attitudes towards material objects. They can also reflect the person's altruism, sense of kindness, or even their religious values.

Now recall that when preferences are complete and transitive, and X is finite the individual can completely order the elements of X from best to worst. The R^i, therefore, convey all the information we need to know to determine the individual's choice from among alternatives in X. To determine the *social* choice, however, we will need some ranking of the social states in X that reflects 'society's' preferences. Ideally, we would like to be able to compare any two alternatives in X from a social point of view, and we would like those binary comparisons to be consistent in the usual way. We have, then, the following definition.

DEFINITION 6.1 ***A Social Preference Relation***

A social preference relation, R, is a complete and transitive binary relation on the set X of social states. For x and y in X, we read xRy as the statement 'x is socially at least as good as y'. We let P and I be the associated relations of strict social preference and social indifference, respectively.

We take it for granted that the ranking of alternatives from a social point of view should depend on how individuals rank them. The problem considered by Arrow can be simply put. How can we go from the often divergent, but individually consistent, personal views of society's members to a single and consistent social view?

This is not an easy problem at all. When we insist on transitivity as a criterion for consistency in social choice, certain well-known difficulties can easily arise. For example, **Condorcet's paradox** illustrates that the familiar method of majority voting can fail to satisfy the transitivity requirement on R. To see this, suppose $N = 3$, $X = \{x, y, z\}$, and

individual (strict) preferences over X are as follows

Person 1	Person 2	Person 3
x	y	z
y	z	x
z	x	y

In a choice between x and y, x would get two votes and y would get one, so the social preference under majority rule would be xPy. In a choice between y and z, majority voting gives yPz. Because xPy and yPz, transitivity of social preferences would require that xPz. However, with these individual preferences, z gets two votes to one for x, so majority voting here would give the social preference as zPx, thus violating transitivity. Note that in this example, the mechanism of majority rule is 'complete' in that it is capable of giving a best alternative in every possible pairwise comparison of alternatives in X. The failure of transitivity, however, means that within this set of three alternatives, no single best alternative can be determined by majority rule. Requiring completeness *and* transitivity of the social preference relation implies that it must be capable of placing every element in X within a hierarchy from best to worst. The kind of consistency required by transitivity has, therefore, considerable structural implications.

Yet consistency, alone, is not particularly interesting or compelling in matters of social choice. One can be perfectly consistent and still violate every moral precept the community might share. The more interesting question to ask might be put like this: how can we go from consistent individual views to a social view that is consistent and that *also* respects certain basic values on matters of social choice that are shared by members of the community? Because disagreement among individuals on matters of 'basic values' is in fact the very reason a problem of *social* choice arises in the first place, we will have to be very careful indeed in specifying these if we want to keep from trivialising the problem at the outset.

With such cautions in mind, however, we can imagine our problem as one of finding a 'rule', or function, capable of aggregating and reconciling the different individual views represented by the individual preference relations R^i into a single social preference relation R satisfying certain ethical principles. Formally, then, we seek a **social welfare function**, f, where

$$R = f(R^1, \ldots, R^N).$$

Thus, f takes an N-tuple of individual preference relations on X and turns (maps) them into a social preference relation on X.

For the remainder of this subsection we shall suppose that the set of social states, X, is finite.

Arrow has proposed a set of four conditions that might be considered minimal properties the social welfare function, f, should possess. They are as follows.

ASSUMPTION 6.1 ***Arrow's Requirements on the Social Welfare Function***

U. Unrestricted Domain. *The domain of f must include all possible combinations of individual preference relations on X.*

WP. Weak Pareto Principle. *For any pair of alternatives x and y in X, if xP^iy for all i, then xPy.*

IIA. Independence of Irrelevant Alternatives. *Let $R = f(R^1, \ldots, R^N)$, $\tilde{R} = f(\tilde{R}^1, \ldots, \tilde{R}^N)$, and let x and y be any two alternatives in X. If each individual i ranks x versus y under R^i the same way that he does under $\tilde{R}^i$, then the social ranking of x versus y is the same under R and $\tilde{R}$.*

D. Non-dictatorship. *There is no individual i such that for all x and y in X, xP^iy implies xPy regardless of the preferences R^j of all other individuals $j \neq i$.*

Condition *U* says that *f* is able to generate a social preference ordering regardless of what the individuals' preference relations happen to be. It formalises the principle that the ability of a mechanism to make social choices should not depend on society's members holding any particular sorts of views. As we have seen before, this condition, together with the transitivity requirement on *R*, rules out majority voting as an appropriate mechanism because it sometimes fails to produce a transitive social ordering when there are more than three alternatives to consider.

Condition *WP* is very straightforward, and one that economists, at least, are quite comfortable with. It says society should prefer *x* to *y* if every single member of society prefers *x* to *y*. Notice that this is a *weak* Pareto requirement because it does not specifically require the social preference to be for *x* if, say, all but one strictly prefer *x* to *y*, yet one person is indifferent between *x* and *y*.

Condition *IIA* is perhaps the trickiest to interpret, so read it over carefully. In brief, the condition says that the social ranking of *x* and *y* should depend only on the individual rankings of *x* and *y*. Note that the individual preferences R^i and $\tilde{R}^i$ are allowed to differ in their rankings over pairs other than *x*, *y*. As you consider for yourself the reasonableness of *IIA*, think of what could happen if we failed to require it. For example, suppose that in the morning, all individuals rank *z* below both *x* and *y*, but some prefer *x* to *y* and others prefer *y* to *x*. Now suppose that given these individual preferences, the social welfare function leads to a social preference of *x* strictly preferred to *y*. So in the morning, if a choice were to be made between *x* and *y*, 'society' would choose *x*. As it happens, however, a choice between *x* and *y* is postponed until the afternoon. But by then, suppose that the individual preferences have changed so that now *z* is ranked *above* both *x* and *y* by all individuals. However, each individual's ranking of *x* versus *y* remains *unchanged*. Would it be reasonable for the social preference to now switch to *y* being ranked above *x*? *IIA* says it would not.

Condition *D* is a very mild restriction indeed. It simply says there should be no single individual who 'gets his way' on *every* single social choice, regardless of the views

of everyone else in society. Thus, only the most extreme and absolute form of dictatorship is specifically excluded. Not even a 'virtual' dictator, one who always gets his way on all but *one* pair of social alternatives, would be ruled out by this condition alone.

Now take a moment to re-examine and reconsider each of these conditions in turn. Play with them, and try to imagine the kind of situations that could arise in a problem of social choice if one or more of them failed to hold. If, in the end, you agree that these are mild and *minimal* requirements for a reasonable social welfare function, you will find the following theorem astounding, and perhaps disturbing.

THEOREM 6.1 ***Arrow's Impossibility Theorem***

If there are at least three social states in X, then there is no social welfare function f that simultaneously satisfies conditions U, WP, IIA, and D.

Proof: The strategy of the proof is to show that conditions *U, WP*, and *IIA* imply the existence of a dictator. Consequently, if *U, WP*, and *IIA* hold, then *D* must fail to hold, and so no social welfare function can satisfy all four conditions.

The proof, following Geanakoplos (1996), proceeds in four steps. Note that axiom *U*, unrestricted domain, is used in each step whenever we choose or alter the preference profile under consideration. Unrestricted domain ensures that every such profile of preferences is admissible.

Step 1: Consider any social state, *c*. Suppose each individual places state *c* at the bottom of his ranking. By WP, the social ranking must place *c* at the bottom as well. See Fig. 6.2.

Step 2: Imagine now moving *c* to the top of individual 1's ranking, leaving the ranking of all other states unchanged. Next, do the same with individual 2: move *c* to the top of 2's ranking. Continue doing this one individual at a time, keeping in mind that each of these changes in individual preferences might have an effect on the social ranking. Eventually, *c* will be at the top of every individual's ranking, and so it must then also be at the top of the social ranking by WP. Consequently, there must be a *first* time during this process that the social ranking of *c* increases. Let individual *n* be the first such that raising *c* to the top of his ranking causes the social ranking of *c* to increase.

R^1	R^2	$\cdots$	R^N	R
x	x'	$\cdots$	x''	x'''
y	y'	$\cdots$	y''	y'''
$\cdot$	$\cdot$		$\cdot$	$\cdot$
$\cdot$	$\cdot$		$\cdot$	$\cdot$
$\cdot$	$\cdot$		$\cdot$	$\cdot$
c	c	$\cdots$	c	c

Figure 6.2. A consequence of *WP* and *U* in the proof of Arrow's theorem.

R^1	R^2	$\cdots$	R^n	$\cdots$	R^N	R
c	c	$\cdots$	c	$\cdots$	x''	c
x	x'	$\cdots$		$\cdots$	y''	
y	y'				$\cdot$	$\cdot$
$\cdot$	$\cdot$				$\cdot$	$\cdot$
$\cdot$	$\cdot$				$\cdot$	$\cdot$
$\cdot$	$\cdot$				$\cdot$	$\cdot$
w	w'	$\cdots$		$\cdots$	c	w'''

Figure 6.3. Axioms *WP*, *U*, and *IIA* yield a pivotal individual.

We claim that, as shown in Fig. 6.3, when c moves to the top of individual n's ranking, the social ranking of c not only increases but c also moves to the *top* of the social ranking.

To see this, assume by way of contradiction that the social ranking of c increases, but not to the top; i.e., $\alpha R c$ and $cR\beta$ for some states $\alpha, \beta \neq c$.

Now, because c is either at the bottom or at the top of every individual's ranking, we can change each individual i's preferences so that $\beta P^i \alpha$, while leaving the position of c unchanged for that individual. But this produces our desired contradiction because, on the one hand, $\beta P^i \alpha$ for every individual implies by WP that β must be strictly preferred to α according to the social ranking; i.e., $\beta P \alpha$. But, on the other hand, because the rankings of c relative to α and of c relative to β have not changed in any individual's ranking, *IIA* implies that the social rankings of c relative to α and of c relative to β must be unchanged; i.e., as initially assumed, we must have $\alpha R c$ and $cR\beta$. But transitivity then implies $\alpha R \beta$, contradicting $\beta P \alpha$. This establishes our claim that c must have moved to the top of the social ranking as in Fig. 6.3.

Step 3: Consider now any two distinct social states a and b, each distinct from c. In Fig. 6.3, change the profile of preferences as follows: change individual n's ranking so that $aP^n cP^n b$, and for every other individual rank a and b in any way so long as the position of c is unchanged for that individual. Note that in the new profile of preferences the ranking of a to c is the same for every individual as it was just *before* raising c to the top of individual n's ranking in Step 2. Therefore, by *IIA*, the social ranking of a and c must be the same as it was at that moment. But this means that aPc because at that moment c was still at the bottom of the social ranking.

Similarly, in the new profile of preferences, the ranking of c to b is the same for every individual as it was just *after* raising c to the top of individual n's ranking in Step 2. Therefore by *IIA*, the social ranking of c and b must be the same as it was at that moment. But this means that cPb because at that moment c had just risen to the top of the social ranking.

So, because aPc and cPb, we may conclude by transitivity that aPb. Note then that no matter how the others rank a and b, the social ranking agrees with individual n's ranking. By *IIA*, and because a and b were arbitrary, we may therefore conclude that for all social

states a and b distinct from c

$$aP^n b \text{ implies } aPb.$$

That is, individual n is a dictator on all pairs of social states not involving c. The final step shows that individual n is in fact a dictator.

Step 4: Let a be distinct from c. We may repeat the above steps with a playing the role of c to conclude that some individual is a dictator on all pairs not involving a. However, recall that individual n's ranking of c (bottom or top) in Fig. 6.3 affects the social ranking of c (bottom or top). Hence, it must be individual n who is the dictator on all pairs not involving a. Because a was an arbitrary state distinct from c, and together with our previous conclusion about individual n, this implies that n is a dictator. ∎

Although here we have cast Arrow's theorem as an 'impossibility' result, the proof just sketched suggests it can also be stated as a 'possibility' result. That is, we have shown that any social welfare function satisfying the three conditions U, WP, and IIA must yield a social preference relation that exactly coincides with one person's preferences whenever that person's preferences are strict. As you are asked to explore in Exercise 6.3 this leaves several 'possibilities' for the social welfare function, although all of them are dictatorial according to condition D.

6.2.1 A DIAGRAMMATIC PROOF

The importance of Arrow's theorem warrants presenting another proof. Our second proof will be diagrammatic, dealing with the case of just two individuals. Together, we hope that the two proofs provide useful insight into the nature of this remarkable result.[1]

We shall depart from the setup of the previous section in several ways. First, we shall assume that X contains not just three or more social states, but infinitely many. Indeed, we assume that X is a non-singleton convex subset of $\mathbb{R}^K$ for some $K \geq 1$.[2]

Second, we assume that the individual preferences R^i on X can be represented by continuous utility functions, $u^i\colon X \to \mathbb{R}$. Thus, our domain of preferences is not completely unrestricted.[3]

Third, we assume that the social welfare function, f, maps profiles of continuous individual utility functions $\mathbf{u}(\cdot) = (u^1(\cdot), \ldots, u^N(\cdot))$ into a *continuous* utility function for society. Therefore, $f(u^1(\cdot), \ldots, u^N(\cdot))$ is a social utility function and $[f(u^1(\cdot), \ldots, u^N(\cdot))](x)$ is the utility assigned to the social state x. Note that the utility assigned to x, namely $[f(u^1(\cdot), \ldots, u^N(\cdot))](x)$, can in principle depend upon each individual's *entire utility function* $u^i(\cdot)$ and not just the utility $u^i(x)$ that each individual assigns to x.

---

[1] The diagrammatic idea of this proof is due to Blackorby, Donaldson, and Weymark (1984).

[2] This assumption can be weakened substantially. For example, the argument we shall provide is valid so long as $X \subseteq \mathbb{R}^K$ contains a point and a sequence of distinct points converging to it.

[3] If X were finite, every R^i would have a utility representation and every utility representation would be continuous. Hence, in the finite case, assuming continuity does not restrict the domain of preferences at all. This is why we assume an infinite X here, so that continuity has 'bite'.

For each continuous $\mathbf{u}(\cdot) = (u^1(\cdot), \ldots, u^N(\cdot))$ we henceforth let $f_{\mathbf{u}}$ denote the social utility function $f(u^1(\cdot), \ldots, u^N(\cdot))$ and we let $f_{\mathbf{u}}(x) = [f(u^1(\cdot), \ldots, u^N(\cdot))](x)$ denote the utility assigned to $x \in X$.

To maintain the idea that the social preference relation is determined only by the individual preference relations, R^i – an idea that is built into the previous section's treatment of Arrow's Theorem – it must be the case that the ordering of the social states according to $f_{\mathbf{u}} = f(u^1(\cdot), \ldots, u^N(\cdot))$ would be unchanged if any $u^i(\cdot)$ were replaced with a utility function representing the same preferences. Thus, because two utility functions represent the same preferences if and only if one is a strictly increasing transformation of the other, the social welfare function f must have the following property: if for each individual i, $u^i\colon X \to \mathbb{R}$ is continuous and $\psi^i\colon \mathbb{R} \to \mathbb{R}$ is strictly increasing and continuous, then

$$f_{\mathbf{u}}(x) \geq f_{\mathbf{u}}(y) \text{ if and only if } f_{\psi \circ \mathbf{u}}(x) \geq f_{\psi \circ \mathbf{u}}(y), \tag{6.1}$$

where $\psi \circ \mathbf{u}(\cdot) = (\psi^1(u^1(\cdot)), \ldots, \psi^N(u^N(\cdot)))$. That is, f must be order-invariant to strictly increasing continuous transformations of individual utility functions, where only continuous transformations ψ^i are considered to ensure that the transformed individual utility functions remain continuous.

Condition *U* in this setup means that the domain of f is the entire set of profiles of continuous individual utility functions. Condition *IIA* means precisely what it meant before, but note in particular it implies that whether $f_{\mathbf{u}}(x)$ is greater, less, or equal to $f_{\mathbf{u}}(y)$ can depend only on the vectors $\mathbf{u}(x) = (u^1(x), \ldots, u^N(x))$ and $\mathbf{u}(y) = (u^1(y), \ldots, u^N(y))$ and not on any other values taken on by the vector function $\mathbf{u}(\cdot) = (u^1(\cdot), \ldots, u^N(\cdot))$.[4] The meanings of conditions *WP* and *D* remain as before.

Consider now imposing the following additional condition on f.

PI. *Pareto Indifference Principle.* *If $u^i(x) = u^i(y)$ for all $i = 1, \ldots, N$, then $f_{\mathbf{u}}(x) = f_{\mathbf{u}}(y)$.*

The Pareto Indifference Principle requires society to be indifferent between two states if each individual is indifferent between them.

It can be shown (see Exercise 6.4 and also Sen (1970a)) that if f satisfies *U*, *IIA*, *WP*, and *PI*, then there is a strictly increasing continuous function, $W\colon \mathbb{R}^N \to \mathbb{R}$, such that for all social states x, y, and every profile of continuous individual utility functions $\mathbf{u}(\cdot) = (u^1(\cdot), \ldots, u^N(\cdot))$,

$$f_{\mathbf{u}}(x) \geq f_{\mathbf{u}}(y) \text{ if and only if } W(u^1(x), \ldots, u^N(x)) \geq W(u^1(y), \ldots, u^N(y)). \tag{6.2}$$

Condition (6.2) says that the social welfare function f can be summarised by a strictly increasing and continuous function W – that we will also call a social welfare function – that simply orders the vectors of individual utility numbers corresponding to

[4] As already noted, the social utility, $f_{\mathbf{u}}(x)$, assigned to the alternative x might depend on each individual's entire utility function. *IIA* goes a long way towards requiring that $f_{\mathbf{u}}(x)$ depend only on the vector of utilities $(u^1(x), \ldots, u^N(x))$.

the alternatives. Consequently, we may restrict our attention to this simpler yet equivalent form of a social welfare function. It is simpler because it states directly that the social utility of an alternative depends only on the vector of individual utilities of that alternative.

Our objective now is to deduce the existence of a dictator from the fact that W satisfies (6.2).

The property expressed in (6.1) that f is order-invariant to continuous strictly increasing transformations of individual utility functions has important implications for the welfare function W. For suppose $(u^1, \ldots, u^N)$ and $(\tilde{u}^1, \ldots, \tilde{u}^N)$ are utility vectors associated with two social states x and y. Combining (6.1) with (6.2) implies that W's ordering of $\mathbb{R}^N$ must be invariant to any continuous strictly increasing transformation of individual utility numbers. Therefore if W ranks x as socially better than y, i.e., if

$$W(u^1, \ldots, u^N) > W(\tilde{u}^1, \ldots, \tilde{u}^N),$$

then we must also have,

$$W(\psi^1(u^1), \ldots, \psi^N(u^N)) > W(\psi^1(\tilde{u}^1), \ldots, \psi^N(\tilde{u}^N))$$

for any N continuous strictly increasing functions, $\psi^i\colon \mathbb{R} \to \mathbb{R}$, $i = 1, 2, \ldots, N$. Appreciating this is key to the argument that follows.

For the diagrammatic proof we assume that $N = 2$ so we can work in the plane.

To begin, consider an arbitrary point $\bar{\mathbf{u}}$ in Fig. 6.4, and try to imagine the social indifference curve on which it lies. For reference, the utility space has been divided into four regions relative to $\bar{\mathbf{u}}$, where the regions do not include the dashed lines. First, note that, by *WP*, all points in region I must be socially preferred to $\bar{\mathbf{u}}$. Similarly, $\bar{\mathbf{u}}$ must be socially preferred to all points in region III. Our problem, then, is to rank points in II, IV, and the excluded boundaries, relative to $\bar{\mathbf{u}}$.

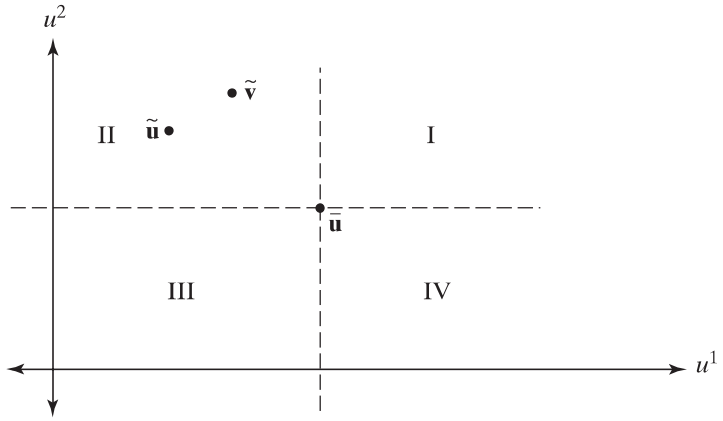


Figure 6.4. A diagrammatic proof of Arrow's theorem.

Now consider an arbitrary point $\tilde{\mathbf{u}}$ in II. One of the following must hold

$$W(\bar{\mathbf{u}}) > W(\tilde{\mathbf{u}}), \tag{6.3}$$

$$W(\bar{\mathbf{u}}) = W(\tilde{\mathbf{u}}), \tag{6.4}$$

$$W(\bar{\mathbf{u}}) < W(\tilde{\mathbf{u}}). \tag{6.5}$$

Suppose for the moment that $W(\bar{\mathbf{u}}) < W(\tilde{\mathbf{u}})$. Then because W's ordering of $\mathbb{R}^N$ is invariant to continuous strictly increasing transformations of utilities, that same ranking must be preserved when we apply *any* continuous strictly increasing transformations to the individuals' utilities. Suppose we choose two strictly increasing functions, ψ^1 and ψ^2, where

$$\psi^1(\bar{u}^1) = \bar{u}^1,$$
$$\psi^2(\bar{u}^2) = \bar{u}^2.$$

Now apply these functions to the coordinates of the point $\tilde{\mathbf{u}}$. Because $\tilde{\mathbf{u}}$ is in region II, we know that $\tilde{u}^1 < \bar{u}^1$ and $\tilde{u}^2 > \bar{u}^2$. Then because the ψ_i are strictly increasing, when applied to $\tilde{\mathbf{u}}$, we must have

$$\tilde{v}^1 \equiv \psi^1(\tilde{u}^1) < \psi^1(\bar{u}^1) = \bar{u}^1, \tag{6.6}$$

$$\tilde{v}^2 \equiv \psi^2(\tilde{u}^2) > \psi^2(\bar{u}^2) = \bar{u}^2. \tag{6.7}$$

Equations (6.6) and (6.7), together, inform us that the point $\tilde{\mathbf{v}} \equiv (\tilde{v}^1, \tilde{v}^2)$ must be somewhere in region II, as well. Because we have complete flexibility in our choice of the continuous strictly increasing ψ^i, we can, by an appropriate choice, map $\tilde{\mathbf{u}}$ into *any* point in region *II*.[5] But then because the social ranking of the underlying social states must be invariant to such transforms of individuals' utility, *every* point in region II must be ranked *the same way* relative to $\bar{\mathbf{u}}$! If, as we supposed, $W(\bar{\mathbf{u}}) < W(\tilde{\mathbf{u}})$, then every point in region II must be preferred to $\bar{\mathbf{u}}$. Yet nowhere in the argument did we use the fact that $W(\bar{\mathbf{u}}) < W(\tilde{\mathbf{u}})$. We could have begun by supposing any of (6.3), (6.4), or (6.5), and reached the same general conclusion by the same argument. Thus, under the invariance requirements on individual utility, *every* point in region II must be ranked in one of three ways relative to $\bar{\mathbf{u}}$: either $\bar{\mathbf{u}}$ is preferred, indifferent to, or worse than every point in region II. We will write this as the requirement that exactly one of the following must hold:

$$W(\bar{\mathbf{u}}) > W(\text{II}), \tag{6.8}$$

$$W(\bar{\mathbf{u}}) = W(\text{II}), \tag{6.9}$$

$$W(\bar{\mathbf{u}}) < W(\text{II}). \tag{6.10}$$

Note that (6.9) certainly cannot hold, for this would mean that all points in region II, being indifferent (under W) to $\bar{\mathbf{u}}$, are indifferent to one another. But this contradicts

[5] For example, to obtain $\psi^i(\bar{u}^i) = \bar{u}^i$ and $\psi^i(\tilde{u}^i) = u^i$ we can choose the continuous function

$$\psi^i(t) \equiv \left[\frac{\bar{u}^i - u^i}{\bar{u}^i - \tilde{u}^i}\right] t + \left[\frac{u^i - \tilde{u}^i}{\bar{u}^i - \tilde{u}^i}\right] \bar{u}^i,$$

which is the form $\psi^i(t) = \alpha^i t + \beta^i$. Note that for any choice of (u^1, u^2) in region II, $\alpha^1, \alpha^2 > 0$.

W being strictly increasing because the point $\tilde{\mathbf{v}} \gg \tilde{\mathbf{u}}$ in region II (see Fig. 6.4) is strictly preferred to $\tilde{\mathbf{u}}$.

So, either $W(\bar{\mathbf{u}}) > W(\text{II})$ or $W(\bar{\mathbf{u}}) < W(\text{II})$. By a parallel argument to the one just given, we could consider points in region IV and show that either $W(\bar{\mathbf{u}}) > W(\text{IV})$ or $W(\bar{\mathbf{u}}) < W(\text{IV})$.

Now, suppose that $W(\bar{\mathbf{u}}) < W(\text{II})$. Then, in particular, $W(\bar{\mathbf{u}}) < W(\bar{u}^1 - 1, \bar{u}^2 + 1)$. Consider the pair of strictly increasing functions $\psi^1(u^1) = u^1 + 1$, $\psi^2(u^2) = u^2 - 1$. Applying these to $\bar{\mathbf{u}}$ and $(\bar{u}^1 - 1, \bar{u}^2 + 1)$ maps them into the points $(\bar{u}^1 + 1, \bar{u}^2 - 1)$ and $\bar{\mathbf{u}}$, respectively. But because W must be order-invariant to such transforms, these images must be ordered in the same way as their inverse images are ordered. Consequently, we must have $W(\bar{u}^1 + 1, \bar{u}^2 - 1) < W(\bar{\mathbf{u}})$. But this means that $\bar{\mathbf{u}}$ is strictly socially preferred to the point $(\bar{u}^1 + 1, \bar{u}^2 - 1)$ in region IV. Consequently, $\bar{\mathbf{u}}$ must be strictly socially preferred to every point in region IV.

So, we have shown that if $W(\bar{\mathbf{u}}) < W(\text{II})$, then $W(\bar{\mathbf{u}}) > W(\text{IV})$. A similar argument establishes that if $W(\bar{\mathbf{u}}) > W(\text{II})$, then $W(\bar{\mathbf{u}}) < W(\text{IV})$. Altogether, we have so far shown that

$$\text{either} \qquad W(\text{IV}) < W(\bar{\mathbf{u}}) < W(\text{II}), \tag{6.11}$$

$$\text{or} \qquad W(\text{II}) < W(\bar{\mathbf{u}}) < W(\text{IV}). \tag{6.12}$$

Now, note that if adjacent regions are ranked the same way relative to $\bar{\mathbf{u}}$, then the dashed line separating the two regions must be ranked that same way relative to $\bar{\mathbf{u}}$. For example, suppose regions I and II are ranked above $\bar{\mathbf{u}}$. Since by *WP* any point on the dashed line above $\bar{\mathbf{u}}$ is ranked above points in region II that lie strictly below it, transitivity implies this point on the dashed line must be ranked above $\bar{\mathbf{u}}$.

Consequently, if (6.11) holds, then because region I is ranked above $\bar{\mathbf{u}}$ and region III is ranked below, the social ranking must be as given in Fig. 6.5(a), where '+' ('−') denotes

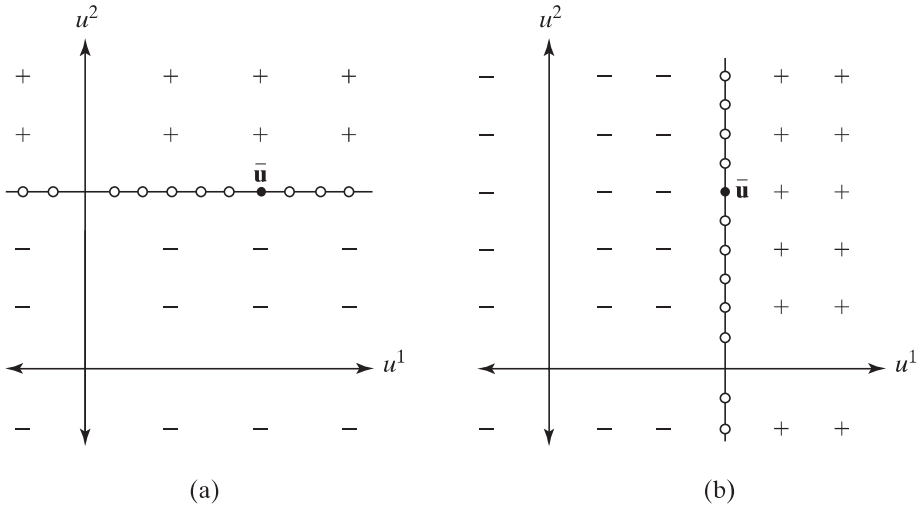


Figure 6.5. Social welfare possibilities under Arrow's conditions.

utility vectors $\mathbf{u} = (u^1, u^2)$ with $W(\mathbf{u})$ greater than (less than) $W(\bar{\mathbf{u}})$. But the continuity of W then implies that the indifference curve through $\bar{\mathbf{u}}$ is a horizontal straight line. On the other hand, if instead (6.12) holds so that Fig. 6.5(b) is relevant, then the indifference curve through $\bar{\mathbf{u}}$ would be a vertical straight line.

So, because $\bar{\mathbf{u}}$ was arbitrary, we may conclude that the indifference curve through every utility vector is either a horizontal or a vertical straight line. However, because indifference curves cannot cross one another, this means that either *all* indifference curves are horizontal straight lines, in which case individual 2 would be a dictator, or *all* indifference curves are vertical straight lines, in which case individual 1 is a dictator. In either case, we have established the existence of a dictator and the proof is complete.

6.3 Measurability, Comparability, and Some Possibilities

Arrow's theorem is truly disturbing. A very careful look at each of his requirements should impress you with their individual reasonableness and their collective economy. Only the very bold can be sanguine about dropping or relaxing any one of them. Yet the import of the theorem is that this is precisely what we must be prepared to do.

There have been various attempts to rescue social welfare analysis from the grip of Arrow's theorem. One has been to relax the requirements that must be satisfied by the social relation R. For example, replacing transitivity of R with a weaker restriction called 'acyclicity', and replacing the requirement that R order all alternatives from best to worse with the simpler restriction that we be merely capable of finding a best alternative among any subset, opens the way to several possible choice mechanisms, each respecting the rest of Arrow's conditions. Similarly, if transitivity is retained, but condition U is replaced with the assumption that individual preferences are 'single-peaked', Black (1948) has shown that majority voting satisfies the rest of Arrow's conditions, provided that the number of individuals is odd!

Another approach has proceeded along different lines and has yielded interesting results. Rather than argue with Arrow's conditions, attention is focused instead on the information assumed to be conveyed by individuals' preferences. In Arrow's framework, only the individuals' preference relations, R^i, are used as data in deriving the social preference relation $R = f(R^1, \ldots, R^N)$. Thus, if a society wants to implement f, it would obtain from each individual his ranking of the states from best to worst. From this data alone f would provide a ranking of the social states. Obviously, this process yields no information whatsoever about the strength of any particular individual's preferences for x in comparison to another individual's preference for y, nor does it yield any information about how much more one individual favours x over y in comparison to how much more another individual favours y over x. By design, Arrow's approach does not consider such information.

The alternative is to think about what would occur if such information were considered. Before merely pushing forward, a warning is in order. The idea that 'intensity of preference' can be compared in a coherent way across individuals is controversial at best. Nonetheless, the alternative approach to social choice that we are about to explore takes as

a starting point – as an assumption – that such comparisons can be made in a meaningful way. We shall not attempt to justify this assumption. Let us just see what it can do for us.

The basic references for this line of work include Hammond (1976), d'Aspremont and Gevers (1977), Roberts (1980), and Sen (1984). Here, we will only consider a few of their findings to try and get the flavour.

To get us started, consider a situation with just two individuals. Suppose that individual 1 prefers state x to y and that individual 2 prefers y to x. In such a symmetric situation, more information might be useful in order to make a social choice. Indeed, suppose for example that society wishes to make its least well off individual as well off as as possible. It would then be useful to know whether individual 1's welfare from the state that he least prefers, namely y, is greater than 2's welfare from the state he least prefers, namely x. Suppose – and here is the important assumption – that the individual utility numbers provide this information. That is, suppose that i's utility function is $u^i(\cdot)$, that $u^1(y)$ is greater than $u^2(x)$, and that this is interpreted to mean that 1 *is better off at* y *than* 2 *is at* x. Armed with the additional information that the least well off individual is better off at y than at x, this society's social welfare function ranks y strictly above x.

Next, suppose that the two individual utility functions are $v^1(\cdot)$ and $v^2(\cdot)$ and that it is still the case that 1 prefers x to y and 2 prefers y to x, but now $v^1(y)$ is less than $v^2(x)$. That is, it is now the case that 1 is worse off at y than 2 is at x. Because the least well off individual is better off at x, this society now strictly prefers x to y *even though the individual rankings over x and y did not change.*

The point of this example is to demonstrate that if utilities carry more meaning than simply the ranking of states, then the social welfare function need not be invariant to strictly increasing utility transformations. The reason is that while strictly increasing transformations preserve utility comparisons between states for each individual separately, they need not preserve utility rankings between states across individuals. To guarantee that $\psi^i(u^i(x)) \geq \psi^j(u^j(y))$ whenever $u^i(x) \geq u^j(y)$, the utility transformations ψ^i and ψ^j must be strictly increasing and identical, i.e., $\psi^i = \psi^j$. Thus, the social welfare function f would need to be invariant only to strictly increasing utility transformations that are identical across individuals. This more limited set of restrictions allows more possibilities for f and a chance to avoid the impossibility result. When a social welfare function f is permitted to depend *only* on the ordering of utilities both for and across individuals, it must be invariant to *arbitrary*, but common, strictly increasing individual utility transformations. We will then say that f is **utility-level invariant**.

A second type of information that might be useful in making social choices is a measure of how much individual i gains when the social state is changed from x to y in comparison to how much individual j loses. In this case it is assumed that individual i's gain in the move from x to y is the difference in his utilities $u^i(y) - u^i(x)$ and that $u^i(y) - u^i(x) \geq u^j(x) - u^j(y)$ means that i's gain is at least as large as j's loss. Again, if a social welfare function is permitted to take such information into account then it need not be invariant to utility transformations that fail to preserve this information. It is not difficult to see that in order to preserve comparisons of utility differences across individuals, each individual i's utility transformation must be of the form $\psi^i(u^i) = a^i + bu^i$, where $b > 0$ is common to all individuals.

When a social welfare function f is permitted to depend *only* on the ordering of utility differences both for and across individuals, it must be invariant to arbitrary strictly increasing individual utility transformations of the form $\psi^i(u^i) = a^i + bu^i$, where $b > 0$. We'll then say that f is **utility-difference invariant**.

Other forms of measurability and interpersonal comparability can be imagined and combined in various ways, but we just stick with the two considered above. For later reference, we summarise the previous discussion as follows, where a social welfare function f maps profiles of utility functions into a social utility function.

DEFINITION 6.2 ***Measurability, Comparability, and Invariance***

1. *A social welfare function f is utility-level invariant if it is invariant to arbitrary, but common, strictly increasing transformations ψ applied to every individual's utility function. Hence, f is permitted to depend only on the ordering of utilities both for and across individuals.*
2. *A social welfare function f is utility-difference invariant if it is invariant to strictly increasing transformations of the form $\psi^i(u^i) = a^i + bu^i$, where $b > 0$ is common to each individual's utility transformation. Hence, f is permitted to depend only on the ordering of utility differences both for and across individuals.*

Throughout the remainder of this section we will assume that the set of social states X is a non-singleton convex subset of Euclidean space and that all social choice functions, f, under consideration satisfy **strict welfarism** (i.e., *U*, *WP*, *IIA*, and *PI*), where *U* means that f maps continuous individual utility functions into a continuous social utility function.[6] Consequently (see (6.2) and Exercise 6.4) we may summarise f with a strictly increasing continuous function $W\colon \mathbb{R}^N \to \mathbb{R}$ with the property that for every continuous $\mathbf{u}(\cdot) = (u^1(\cdot), \ldots, u^N(\cdot))$ and every pair of states x and y,

$$f_{\mathbf{u}}(x) \geq f_{\mathbf{u}}(y) \text{ if and only if } W(u^1(x), \ldots, u^N(x)) \geq W(u^1(y), \ldots, u^N(y)),$$

where we remind the reader that $f_{\mathbf{u}}(x)$ is the social utility assigned to x when the profile of individual utility functions is $\mathbf{u}(\cdot) = (u^1(\cdot), \ldots, u^N(\cdot))$.

The extent to which utility is assumed to be measurable and interpersonally comparable can best be viewed as a question of how much information society uses when making social decisions. This is quite distinct from the kind of ethical restrictions a society might wish those decisions to respect. There is, of course, some ethical content to the conditions *U*, *WP*, *IIA* and *PI* embodied in strict welfarism. However, a society may be willing to go further and build even more ethical values into its social welfare function. Each amounts to imposing an extra requirement on the strictly increasing and continuous social welfare function, W. Here, we consider only two.

[6] Sen (1970a) defines f to satisfy *welfarism* if f satisfies *U*, *IIA*, and *PI*.

DEFINITION 6.3 ***Two More Ethical Assumptions on the Social Welfare Function***

A. *Anonymity. Let* $\bar{\mathbf{u}}$ *be a utility N-vector, and let* $\tilde{\mathbf{u}}$ *be another vector obtained from* $\bar{\mathbf{u}}$ *after some permutation of its elements. Then* $W(\bar{\mathbf{u}}) = W(\tilde{\mathbf{u}})$.

HE. *Hammond Equity. Let* $\bar{\mathbf{u}}$ *and* $\tilde{\mathbf{u}}$ *be two distinct utility N-vectors and suppose that* $\bar{u}^k = \tilde{u}^k$ *for all k except i and j. If* $\bar{u}^i < \tilde{u}^i < \tilde{u}^j < \bar{u}^j$, *then* $W(\tilde{\mathbf{u}}) \geq W(\bar{\mathbf{u}})$.

Condition *A* simply says people should be treated symmetrically. Under *A*, the ranking of social states should not depend on the identity of the individuals involved, only the levels of welfare involved. Condition *HE* is slightly more controversial. It expresses the idea that society has a preference towards decreasing the dispersion of utilities across individuals. (Note that there is less dispersion of utilities under $\bar{\mathbf{u}}$ than under $\tilde{\mathbf{u}}$. Nevertheless, can you think of why one might object to ranking $\bar{\mathbf{u}}$ above $\tilde{\mathbf{u}}$?) In what follows, we use these conditions to illustrate how some well-known social welfare functions can be characterised axiomatically.

6.3.1 THE RAWLSIAN FORM

In the ethical system proposed by Rawls (1971), the welfare of society's worst-off member guides social decision making. In the following theorem, we give an axiomatic characterisation of this criterion of social welfare. The proof we provide is diagrammatic and so again we restrict ourselves to the case of $N = 2$.[7]

THEOREM 6.2 ***Rawlsian Social Welfare Functions***

A strictly increasing and continuous social welfare function W satisfies HE if and only if it can take the Rawlsian form, $W = \min[u^1, \ldots, u^N]$. *Moreover, W then satisfies A and is utility-level invariant.*

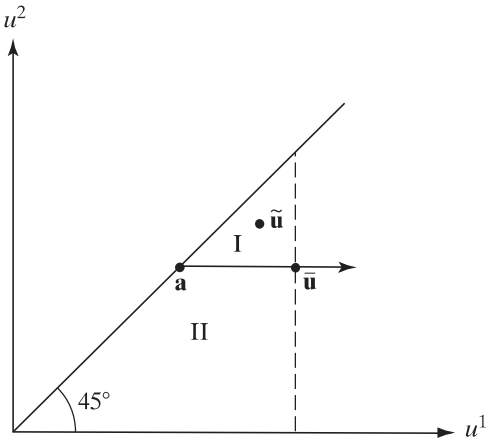
Proof: Suppose that W is continuous, strictly increasing and satisfies *HE*. We must show that it can take the form $W = \min[u^1, \ldots, u^N]$, i.e., that $W(\bar{\mathbf{u}}) \geq W(\tilde{\mathbf{u}})$ if and only if $\min[\bar{u}^1, \ldots, \bar{u}^N] \geq \min[\tilde{u}^1, \ldots, \tilde{u}^N]$.

We prove this diagrammatically only for $N = 2$ by once again characterising the map of social indifference curves. Consult Fig. 6.6 throughout the proof. To begin, choose an arbitrary point **a** on the 45° line and consider the infinite ray extending from **a** to the right. We shall first argue that every point on this ray is socially indifferent to **a** according to W.

Consider an arbitrary point $\bar{\mathbf{u}} = (\bar{u}^1, \bar{u}^2)$ on the ray. We wish to show that $W(\bar{\mathbf{u}}) = W(\mathbf{a})$. Let region *I* denote the region to the left of $\bar{\mathbf{u}}$ below the 45° and above the ray, and let region *II* denote the region to the left of $\bar{\mathbf{u}}$ below the 45° line and below the ray. Thus the ray is in neither region. Consider now an arbitrary point $\tilde{\mathbf{u}} = (\tilde{u}^1, \tilde{u}^2)$ in region *I*. One can easily see that to be in *I*, $\tilde{\mathbf{u}}$ must satisfy the inequalities $\bar{u}^2 < \tilde{u}^2 < \tilde{u}^1 < \bar{u}^1$. (Think

[7] For $N > 2$, see Exercise 6.8 and also Hammond (1976).

Figure 6.6. Proof of Theorem 6.2.



about this.) But then *HE* implies that $W(\tilde{\mathbf{u}}) \geq W(\bar{\mathbf{u}})$. Since $\tilde{\mathbf{u}}$ was an arbitrary point in I, the social utility of every point in I is at least $W(\bar{\mathbf{u}})$, which we write as $W(I) \geq W(\bar{\mathbf{u}})$.[8] As for region II, we must have $W(II) < W(\bar{\mathbf{u}})$ because every point in region II is south-west of $\bar{\mathbf{u}}$ and W is strictly increasing. Thus, we have shown that,

$$W(I) \geq W(\bar{\mathbf{u}}) > W(II). \tag{P.1}$$

Notice now that for every point on the line joining $\mathbf{a}$ and $\bar{\mathbf{u}}$ there are arbitrarily nearby points in region I each of which we have shown to receive social utility at least $W(\bar{\mathbf{u}})$ and there are arbitrarily nearby points in region II each of which we have shown to receive social utility less than $W(\bar{\mathbf{u}})$. Hence, by the continuity of W, every point on the line joining $\mathbf{a}$ and $\bar{\mathbf{u}}$ must receive social utility equal to $W(\bar{\mathbf{u}})$. In particular, $W(\mathbf{a}) = W(\bar{\mathbf{u}})$, as we wished to show. Because $\bar{\mathbf{u}}$ was an arbitrary point on the infinite ray starting at $\mathbf{a}$ and extending rightwards, we conclude that every point on this ray is socially indifferent to $\mathbf{a}$.

An analogous argument to that just given shows also that every point on the infinite ray starting at $\mathbf{a}$ and extending upwards is also socially indifferent to $\mathbf{a}$. Because W is strictly increasing, no other points can be indifferent to $\mathbf{a}$ and therefore the union of these two rays is the social indifference curve through $\mathbf{a}$. Because $\mathbf{a}$ was an arbitrary point on the 45° line, the social indifference map for W is therefore as shown in Fig. 6.7, with indifference curves further from the origin receiving higher social utility because W is strictly increasing. Thus W has the same indifference map as the function $\min[u^1, u^2]$, as desired.

Finally, we note that if $W = \min[u^1, \ldots, u^N]$ then A and HE are easily shown to be satisfied. Moreover, if $\psi : \mathbb{R} \to \mathbb{R}$ is strictly increasing, then $W(\psi(u^1), \ldots, \psi(u^N)) = \psi(W(u^1, \ldots, u^N))$ and therefore $W(\psi(u^1), \ldots, \psi(u^N)) \geq W(\psi(\tilde{u}^1), \ldots, \psi(\tilde{u}^N))$ if and only if $W(u^1, \ldots, u^N) \geq W(\tilde{u}^1, \ldots, \tilde{u}^N)$. Hence, W is utility-level invariant. ∎

[8] In fact, $W(I) > W(\bar{\mathbf{u}})$ because $N = 2$ and W is strictly increasing, but we will not need the strict inequality.

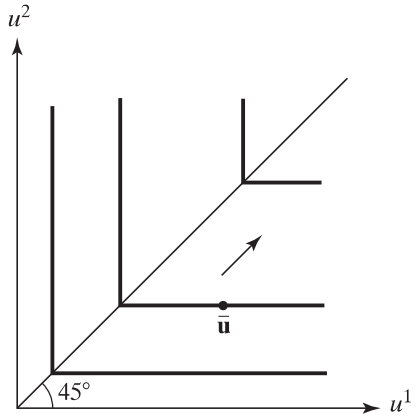


Figure 6.7. Social indifference curves for the (Rawlsian) social welfare function.

6.3.2 THE UTILITARIAN FORM

The utilitarian form is by far the most common and widely applied social welfare function in economics. Under a utilitarian rule, social states are ranked according to the linear sum of utilities. When ranking two social states, therefore, it is the linear sum of the individual utility differences between the states that is the determining factor. Consequently, statements of the form 'in the move from x to y, individual 1 gains more than individual 2' must be meaningful. Thus, utility differences must be comparable both for and across individuals and so we expect the utilitarian social choice function to be related to the property of utility-difference invariance. The theorem to follow shows that this is indeed the case. Once again, our proof covers the $N = 2$ case, the extension to $N > 2$ being straightforward.

THEOREM 6.3 ***Utilitarian Social Welfare Functions***

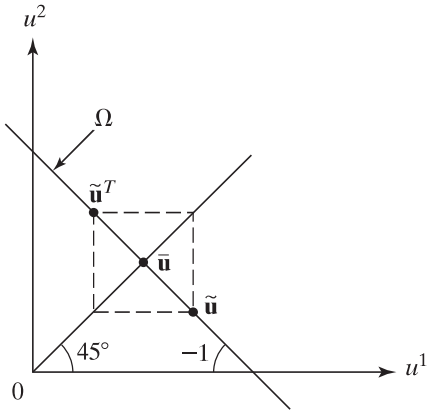
A strictly increasing and continuous social welfare function W satisfies A and utility-difference invariance if and only if it can take the utilitarian form, $W = \sum_{i=1}^{n} u^i$.

Proof: It is clear that if $W = \sum_{i=1}^{N} u^i$, then the conditions of the theorem are satisfied. It remains to show the converse. We will give a diagrammatic proof for the two-person case, but this can be extended to any number of individuals.

In Fig. 6.8, choose any point $\bar{\mathbf{u}} = (\bar{u}^1, \bar{u}^2)$ lying along the 45° line. Define the constant, $\gamma \equiv \bar{u}^1 + \bar{u}^2$ and consider the set of points $\Omega \equiv \{(u^1, u^2) \mid u^1 + u^2 = \gamma\}$. These are all the points lying on a straight line through $\bar{\mathbf{u}}$ with a slope of -1. Choose any point in Ω, distinct from $\bar{\mathbf{u}}$, such as $\tilde{\mathbf{u}}$. Point $\tilde{\mathbf{u}}^T$ is obtained by permuting the element of $\tilde{\mathbf{u}}$, and so $\tilde{\mathbf{u}}^T = (\tilde{u}^2, \tilde{u}^1)$ must also be in Ω. By condition A, $\tilde{\mathbf{u}}$ and $\tilde{\mathbf{u}}^T$ must be ranked the same way relative to $\bar{\mathbf{u}}$.

Now suppose that $W(\bar{\mathbf{u}}) > W(\tilde{\mathbf{u}})$. Under utility-difference dependence, this ranking must be invariant to transformations of the form $\alpha^i + bu^i$. Let $\psi^i(u^i) \equiv (\bar{u}^i - \tilde{u}^i) + u^i$, for $i = 1, 2$. Note carefully that both of these are in the allowable form. Taking note that $2\bar{u}^i = \tilde{u}^1 + \tilde{u}^2$ because $\bar{\mathbf{u}}$ is on the 45° line and both $\bar{\mathbf{u}}$ and $\tilde{\mathbf{u}}$ are in Ω, we apply

Figure 6.8. The utilitarian social welfare function.



these transforms to $\tilde{\mathbf{u}}$ and obtain $(\psi^1(\tilde{u}^1), \psi^2(\tilde{u}^2)) = \bar{\mathbf{u}}$, and apply them to $\bar{\mathbf{u}}$ to obtain $(\psi^1(\bar{u}^1), \psi^2(\bar{u}^2)) = \tilde{\mathbf{u}}^T$. So, these transforms map $\tilde{\mathbf{u}}$ into $\bar{\mathbf{u}}$ and map $\bar{\mathbf{u}}$ into $\tilde{\mathbf{u}}^T$. Thus, if $W(\bar{\mathbf{u}}) > W(\tilde{\mathbf{u}})$, as we have assumed, then by the invariance requirement, we must likewise have $W(\tilde{\mathbf{u}}^T) > W(\bar{\mathbf{u}})$. But together these imply $W(\tilde{\mathbf{u}}^T) > W(\tilde{\mathbf{u}})$, violating *A*, so $W(\bar{\mathbf{u}}) > W(\tilde{\mathbf{u}})$ cannot hold. If, instead, we suppose $W(\tilde{\mathbf{u}}) > W(\bar{\mathbf{u}})$, then by using a similar argument, we get a similar contradiction. We therefore conclude that $W(\bar{\mathbf{u}}) = W(\tilde{\mathbf{u}})$. Condition *A* then tells us $W(\tilde{\mathbf{u}}^T) = W(\bar{\mathbf{u}}) = W(\tilde{\mathbf{u}})$. Now recall that $\tilde{\mathbf{u}}$ was chosen arbitrarily in Ω, so the same argument can be made for any point in that set, and so we have $W(\Omega) = W(\bar{\mathbf{u}})$.

Because W is strictly increasing, every point north-east of Ω must be strictly preferred to every point in Ω, and every point south-west must be strictly worse. Thus, Ω is indeed a social indifference curve, and the social indifference map is a set of parallel straight lines, each with a slope of -1, with social preference increasing north-easterly. This, of course, implies the social welfare function can be chosen to be of the form $W = u^1 + u^2$, completing the proof. ∎

If we drop the requirement of anonymity, the full range of *generalised utilitarian* orderings is allowed. These are represented by linear social welfare functions of the form $W = \sum_i a^i u^i$, where $a^i \geq 0$ for all i and $a^j > 0$ for some j. Under generalised utilitarian criteria, the welfare sum is again the important issue, but the welfare of different individuals can be given different 'weight' in the social assessment.

6.3.3 FLEXIBLE FORMS

To some extent, the greater the measurability and comparability of utility, the greater the range of social welfare functions allowed. For example, suppose that the social welfare function can depend upon the ordering of *percentage* changes in utility both for and across individuals, i.e., that information such as 'in going from x to y, the percentage increase in i's utility is greater than the percentage loss in j's utility', namely,

$$\frac{u^i(x) - u^i(y)}{u^i(x)} > \frac{u^j(x) - u^j(y)}{u^j(x)}$$

matters. Then the social welfare function need not be invariant to strictly increasing transformations unless they are identical and linear, (i.e., $\psi(u^i) = bu^i$, where $b > 0$ is common to all individuals) because only these are guaranteed to maintain the ordering of percentage changes in utility both for and across individuals. If the social welfare function f is permitted to depend *only* on the ordering of percentage changes in utility for and across individuals, then it must be invariant to *arbitrary*, but common, strictly increasing individual transformations of utility of the form $\psi(u^i) = bu^i$, where $b > 0$ is common to all individuals and we will then say that f is **utility-percentage invariant**.

Consequently, both the Rawlsian and utilitarian social welfare functions are permitted here. Indeed, a whole class of social welfare functions are now admitted as possibilities. When a continuous social welfare function satisfies strict welfarism, and is invariant to identical positive linear transformations of utilities, social indifference curves must be negatively sloped and radially parallel.

To see this, consider Fig. 6.9. First, choose an arbitrary point $\bar{\mathbf{u}}$. Clearly, as in the example sketched, the social indifference curve through $\bar{\mathbf{u}}$ must be negatively sloped because, by strict welfarism, W is strictly increasing. Now choose any other point on the ray OA through $\bar{\mathbf{u}}$. This point must be of the form $b\bar{\mathbf{u}}$ for some constant $b > 0$. Now choose any other point $\tilde{\mathbf{u}}$ such that $W(\bar{\mathbf{u}}) = W(\tilde{\mathbf{u}})$. By the invariance requirement, we must also have $W(b\bar{\mathbf{u}}) = W(b\tilde{\mathbf{u}})$, where $\tilde{\mathbf{u}}$ and $b\tilde{\mathbf{u}}$ are on the ray OB, as indicated.

We want to show that the slope of the tangent to the social indifference curve at $\bar{\mathbf{u}}$ is equal to the slope of the tangent at $b\bar{\mathbf{u}}$. First, note that the slope of the chord CC approximates the slope of the tangent at $\bar{\mathbf{u}}$, and the slope of the chord DD approximates the slope of the tangent at $b\bar{\mathbf{u}}$. Because the triangles OCC and ODD are similar, the slope of CC is equal to the slope of DD. Now imagine choosing our point $\tilde{\mathbf{u}}$ closer and closer to $\bar{\mathbf{u}}$ along the social indifference curve through $\bar{\mathbf{u}}$. As $\tilde{\mathbf{u}}$ approaches $\bar{\mathbf{u}}$, correspondingly $b\tilde{\mathbf{u}}$ approaches $b\bar{\mathbf{u}}$ along the social indifference curve through $b\bar{\mathbf{u}}$, and the chords CC and DD remain equal in slope. In the limit, the slope of CC converges to the slope of the tangent at $\bar{\mathbf{u}}$, and the slope of DD converges to the slope of the tangent at $b\bar{\mathbf{u}}$. Thus, the slope of the

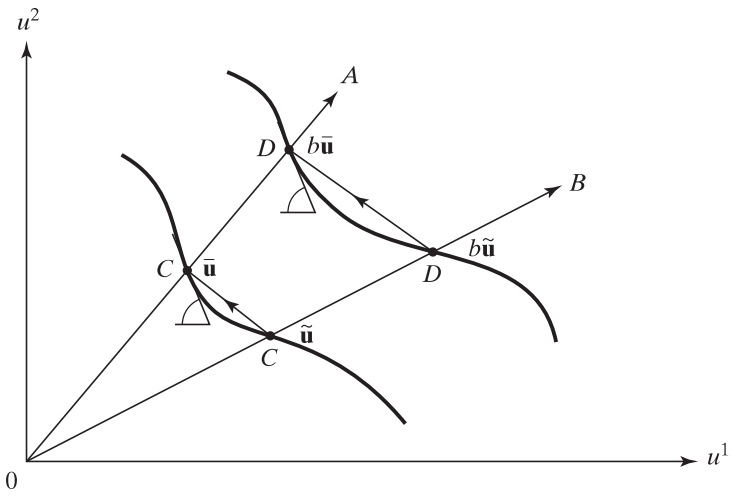


Figure 6.9. Radially parallel social indifference curves.

social indifference curve at $\bar{\mathbf{u}}$ must be equal to the slope of the curve at $b\bar{\mathbf{u}}$. Because $\bar{\mathbf{u}}$ and $b > 0$ were arbitrarily chosen, the slope of every social indifference curve must be the same at every point along a given ray, though, of course, slopes can differ across different rays.

A function's level curves will be radially parallel in this way if and only if the function is *homothetic*. Thus, strict welfarism and utility-percentage invariance allow any continuous, strictly increasing, homothetic social welfare function. If condition *A* is added, the function must be symmetric, and so its social indifference curves must be 'mirror images' around the 45° line. Sometimes a convexity assumption is also added. When the social welfare function is quasiconcave the 'socially at least as good as' sets are convex, and the ethical implication is that inequality in the distribution of welfare, *per se*, is not socially valued. Under strict quasiconcavity, there is a strict bias in favour of equality. (Do you see why?)

Because every homothetic function becomes a linear homogeneous function under some positive monotonic transform, for simplicity let us think in terms of linear homogeneous forms alone. Finally, suppose in addition to *WP*, *A*, and convexity, we add the *strong separability* requirement that the marginal rate of (social) substitution between any two individuals is independent of the welfare of all other individuals. Then the social welfare function must be a member of the CES family:

$$W = \left(\sum_{i=1}^{N} (u^i)^{\rho} \right)^{1/\rho}, \tag{6.13}$$

where $0 \neq \rho < 1$, and $\sigma = 1/(1-\rho)$ is the (constant and equal) elasticity of social substitution between any two individuals.

This is a very flexible social welfare function. Different values for ρ give different degrees of 'curvature' to the social indifference curves, and therefore build in different degrees to which equality is valued in the distribution of welfare. Indeed, the utilitarian form – which implies complete social indifference to how welfare is distributed – can be seen as a limiting case of (6.13) as $\rho \to 1$ ($\sigma \to \infty$). As $\rho \to -\infty$ ($\sigma \to 0$), (6.13) approaches the Rawlsian form, where the social bias in favour of equality is absolute. The range of possibilities is illustrated in Fig. 6.10.

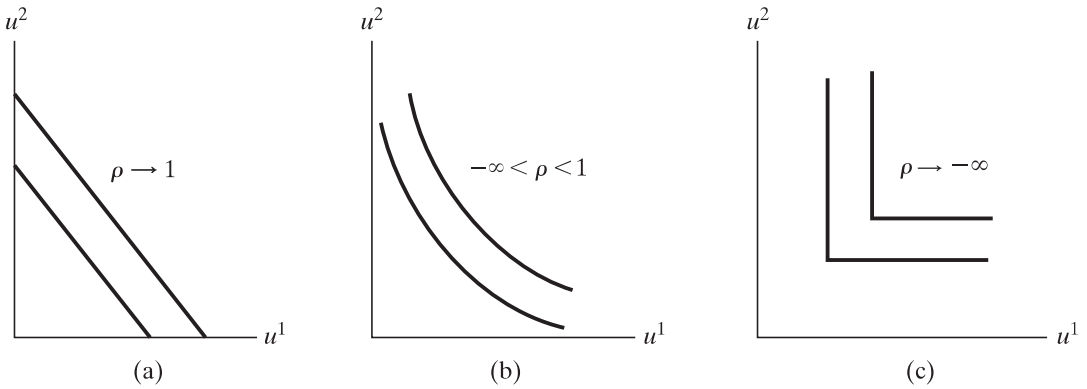


Figure 6.10. CES social welfare functions.

6.4 JUSTICE

Beyond the technical question of what must be assumed in the way of measurability and comparability of utility to sensibly apply a given social welfare function, there is the basic reality that the choice among such functions is effectively a choice between alternative sets of ethical values. On this score, then, matters of opinion really *are* involved. They rightfully belong in the very first stage of any analysis aimed at assessing the *social* significance of economic policies or institutions, when the choice of social welfare function is made.

The literature in economics and the literature in philosophy – one and the same in the days before Adam Smith – have combined again more recently to jointly consider the moral character of the choice that must be made. Guidance has been sought by appeal to axiomatic *theories of justice* that accept the social welfare approach to social decision making. Two broad historical traditions on these questions can be distinguished. One is the utilitarian tradition, associated with Hume, Smith, Bentham, and Mill. The other is the 'contractarian' tradition, associated with Locke, Rousseau, and Kant. More recently, these two traditions have been refined and articulated through the work of Harsanyi (1953, 1955, 1975) and Rawls (1971), respectively.

Both Harsanyi and Rawls accept the notion that a 'just' criterion of social welfare must be one that a rational person would choose if he were 'fair-minded'. To help ensure that the choice be fair-minded, each imagines an 'original position', behind what Rawls calls a 'veil of ignorance', in which the individual contemplates this choice without knowing what his personal situation and circumstances in society actually will be. Thus, each imagines the kind of choice to be made as a choice under uncertainty over who you will end up having to be in the society you prescribe. The two differ, however, in what they see as the appropriate decision rule to guide the choice in the original position.

Harsanyi's approach is remarkably straightforward. First, he accepts the von Neumann-Morgenstern axiomatic description of rationality under conditions of uncertainty. Thus, a person's preferences can be represented by a VNM utility function over social states, $u^i(x)$, which is unique up to positive affine transforms. By the **principle of insufficient reason**, he then suggests that a rational person in the original position must assign an equal probability to the prospect of being in any other person's shoes within the society. If there are N people in society, there is therefore a probability $1/N$ that i will end up in the circumstances of any other person j. Person i therefore must imagine those circumstances and imagine what his preferences, $u^j(x)$, would be. Because a person might end up with any of N possible 'identities', a 'rational' evaluation of social state x then would be made according to its *expected utility*:

$$\sum_{i=1}^{N} (1/N)\, u^i(x). \tag{6.14}$$

In a social choice between x and y, the one with the higher expected utility in (6.14) must be preferred. But this is equivalent to saying that x is socially preferred to y if and only if

$$\sum_{i=1}^{N} u^i(x) > \sum_{i=1}^{N} u^i(y),$$

a purely *utilitarian* criterion.

Rawls rejects Harsanyi's utilitarian rule for several reasons. Among them, he objects to the assignment of *any* probability to the prospect of being any particular individual because in the original position, there can be no empirical basis for assigning such probabilities, whether equal or not. Thus, the very notion of choice guided by expected utility is rejected by Rawls. Instead, he views the choice problem in the original position as one under *complete ignorance*. Assuming people are risk averse, he argues that in total ignorance, a rational person would order social states according to how he or she would view them were they to end up as society's worst-off member. Thus, x will be preferred to y as

$$\min[u^1(x), \ldots, u^N(x)] > \min[u^1(y), \ldots, u^N(y)], \tag{6.15}$$

a purely *maximin* criterion.

Ultimately, then, Rawls' own argument for the maximin over the utilitarian rests on the view that people are risk averse. But this cannot be a wholly persuasive argument, as Arrow (1973) has pointed out. For one thing, the VNM utility functions in Harsanyi's construction can be thought to embody any degree of risk aversion whatsoever. Thus, in Harsanyi's framework, nothing precludes individuals from being risk averse in the original position. Moreover, one need not reject the expected utility rule as a basis for choice to arrive at Rawls' criterion.

To see this, take any utility function $u^i(x)$ over social states with certainty. These same preferences, of course, can be represented equally well by the positive monotonic transform, $v^i(x) \equiv -u^i(x)^{-a}$, where $a > 0$. Now suppose $v^i(x)$ is i's VNM utility function over *uncertain* prospects. It is easy to convince yourself that the degree of risk aversion displayed by $v(x)$ is increasing in the parameter a. Now suppose, as Harsanyi does, (1) equal probabilities of having any identity, (2) an ordering of social states according to their expected utility, and so (3) a social welfare function

$$W = \sum_{i=1}^{N} v^i(x) \equiv -\sum_{i=1}^{N} u^i(x)^{-a}. \tag{6.16}$$

Because the ordering of states given by (6.16) has only ordinal significance, it will be exactly the same under the positive monotonic transform of W given by

$$W^* = (-W)^{-1/a} \equiv \left(\sum_{i=1}^{N} u^i(x)^{-a} \right)^{-1/a} \tag{6.17}$$

For $\rho \equiv -a < 0$, this is in the form of (6.11). We have already noted that as $\rho \to -\infty$ ($a \to \infty$), this approaches the maximin criterion (6.13) as a limiting case. Thus, Rawls' maximin criterion – far from being incompatible with Harsanyi's utilitarianism – instead can be seen as a very special case of it, namely, the one that arises when individuals are *infinitely* risk averse.

On reflection, this makes a good deal of sense. Maximin decision rules are appealing in strategic situations where the interests of some rational and fully informed opponent are diametrically opposed to your own. In the kind of thought experiment required in

the original position, there is little obvious justification for adopting such a decision rule, unless, of course, you are extremely (irrationally?) pessimistic.

Once again, your choice of social welfare function is a choice of distributional values and, therefore, a choice of ethical system. The choice is yours.

6.5 Social Choice and the Gibbard-Satterthwaite Theorem

Up to this point in our analysis of the problem of social welfare, we have focused solely on the task of aggregating the preferences of many individuals into a single preference relation for society. This task, as we have seen, is a formidable one. Indeed, it cannot be carried out if we insist on all of Arrow's conditions.

Implicit in our analysis has been the assumption that the true preferences of each individual can be obtained and that society's preferences are then determined according to its social welfare function. But how, exactly, does society find out the preferences of its individual members? One possibility, of course, is to simply ask each individual to report his ranking of the social states. But this introduces a serious difficulty. Individuals would be better off lying about their preferences than reporting them truthfully if a false report leads to a better social state for them.[9] Thus, in addition to the problem of coherently aggregating individual rankings into a social ranking, there is the problem of finding out individual preferences in the first place. The purpose of this section is to address this latter issue head on.

Throughout this section the set of social states, X, is finite and each of the N individuals in society is permitted to have any preference relation at all on X. Thus, we are assuming unrestricted domain, U. Because the purpose of a social ranking of the states in X is presumably to allow society to make a choice from X, let us focus on that choice directly. Specifically, for each profile of individual rankings $\mathbf{R} = (R^1, \ldots, R^N)$, let $c(\mathbf{R}) \in X$ denote society's *choice* from X. We will assume that the range of $c(\cdot)$ is all of X. That is, for every social state $x \in X$ there is some profile of preferences $\mathbf{R}$ such that $c(\mathbf{R}) = x$. Otherwise, we could just as well eliminate the social state x from the set X. Any function $c(\cdot)$ mapping all profiles of individual preferences on X into a choice from X, and whose range is all of X is called a **social choice function**.[10]

Once again, we would like to avoid dictatorship and in the context of social choice functions a dictatorship is defined in the following natural way.

DEFINITION 6.4 *Dictatorial Social Choice Function*

A social choice function $c(\cdot)$ is dictatorial if there is an individual i such that whenever $c(R^1, \ldots, R^N) = x$ it is the case that xR^iy for every $y \in X$.

[9] Another possibility is to attempt to infer an individual's preferences from his observed choice behaviour. But this too is problematic since an individual can alter his choice behaviour to profitably portray to society false preferences.

[10] Not all treatments of this topic include the full range condition in the definition of a social choice function, choosing instead to add the range condition separately. The present treatment is more convenient for our purposes.

Fix for the moment the preference profile, $\mathbf{R}^{-i}$, of all individuals but i and consider two possible preferences, R^i and $\tilde{R}^i$, for individual i. Let $c(R^i, \mathbf{R}^{-i}) = x$ and $c(\tilde{R}^i, \mathbf{R}^{-i}) = y$. Altogether then, we have a situation in which, when the others report the profile $\mathbf{R}^{-i}$, individual i, by choosing to report either R^i or $\tilde{R}^i$ can choose to make the social state either x or y. When would individual i have an incentive to lie about his preferences? Well, suppose his true preferences happen to be R^i and that given these preferences he strictly prefers y to x. If he reports honestly, the social state will be x. But if he lies and instead reports $\tilde{R}^i$, the social state will be y, a choice he strictly prefers. Hence, in this case, he has an incentive to misreport his preferences.

What property would a social choice function have to have so that under no circumstance would any individual have an incentive to misreport his preferences? It must have the following property called **strategy-proofness.**

DEFINITION 6.5 ***Strategy-Proof Social Choice Function***

A social choice function $c(\cdot)$ is strategy-proof when, for every individual, i, for every pair R^i and $\tilde{R}^i$ of his preferences, and for every profile $\mathbf{R}^{-i}$ of others' preferences, if $c(R^i, \mathbf{R}^{-i}) = x$ and $c(\tilde{R}^i, \mathbf{R}^{-i}) = y$, then xR^iy.

Definition 6.5 rules out exactly the situation described above and, with a little thought, you will convince yourself that if a social choice function is strategy-proof, no individual, no matter what his preferences might be, can ever strictly gain by misreporting his preferences no matter what the others report – even if the others lie about their preferences. Conversely, if a social choice function is not strategy-proof, then there is at least one circumstance (and perhaps many) under which some individual can strictly gain by misreporting his preferences.

Thus, requiring a social choice function to be strategy-proof ensures that it is optimal for individuals to report their preferences honestly and so society's choice will be based upon the true preferences of its individual members. Unfortunately, strategy-proofness has deep consequences. Indeed, reminiscent of Arrow's theorem we have another remarkable, though again negative, result due independently to Gibbard (1973) and Satterthwaite (1975).

THEOREM 6.4 ***The Gibbard-Satterthwaite Theorem***

If there are at least three social states, then every strategy-proof social choice function is dictatorial.

Our proof of Theorem 6.4 follows Reny (2001) and is broken into two parts.[11] Part I shows that a strategy-proof social choice function must exhibit two properties – Pareto-efficiency and monotonicity. Part II shows that any monotonic and Pareto-efficient social choice function is dictatorial. To prepare for the proof, we must first define **Pareto-efficient social choice functions** and **monotonic social choice functions**.

[11] In fact, because the full range condition in Reny (2001) is applied to the smaller domain of strict rankings, our Theorem 6.4 is a slightly stronger result. (At least on the face of it; see Exercise 6.22.)

DEFINITION 6.6 ***Pareto-Efficient Social Choice Function***

A social choice function $c(\cdot)$ is Pareto efficient if $c(R^1, \ldots, R^N) = x$ whenever xP^iy for every individual i and every $y \in X$ distinct from x.

Thus, a social choice function is Pareto efficient if whenever x is at the top of every individual's ranking, the social choice is x.

DEFINITION 6.7 ***Monotonic Social Choice Function***

A social choice function $c(\cdot)$ is monotonic if $c(R^1, \ldots, R^N) = x$ implies $c(\tilde{R}^1, \ldots, \tilde{R}^N) = x$ whenever for each individual i and every $y \in X$ distinct from x, $xR^iy \Longrightarrow x\tilde{P}^iy$.

Monotonicity says that the social choice does not change when individual preferences change so that every individual strictly prefers the social choice to any distinct social state that it was originally at least as good as. Loosely speaking, monotonicity says that the social choice does not change when the social choice rises in each individual's ranking. Notice that the individual rankings between pairs of social states other than the social choice are permitted to change arbitrarily.

We are now prepared to prove Theorem 6.4, but one more word before we do. We are *not* assuming either Pareto efficiency or monotonicity. Part 1 of our proof will *prove* that strategy-proofness implies Pareto efficiency and monotonicity. The *only* assumption Theorem 6.4 makes about the social choice function is that it is strategy-proof.

Proof: Suppose that X contains at least three social states and that $c(\cdot)$ is a strategy-proof social choice function. We must show that $c(\cdot)$ is dictatorial. To do so, we break the proof into two parts.

Part 1. *Strategy-proofness implies monotonicity and Pareto efficiency.*[12]

(a) *Monotonicity.* Let $(R^1, \ldots, R^N)$ be an arbitrary preference profile and suppose that $c(R^1, \ldots, R^N) = x$. Fix an individual, i say, and let $\tilde{R}^i$ be a preference for i such that for every $y \in X$ distinct from x, $xR^iy \Longrightarrow x\tilde{P}^iy$. We shall show that $c(\tilde{R}^i, \mathbf{R}^{-i}) = x$.

Suppose, by way of contradiction, that $c(\tilde{R}^i, \mathbf{R}^{-i}) = y \neq x$. Then, given that the others report $\mathbf{R}^{-i}$, individual i, when his preferences are R^i can report truthfully and obtain the social state x or he can lie by reporting $\tilde{R}^i$ and obtain the social state y. Strategy-proofness requires that lying cannot be strictly better than telling the truth. Hence we must have xR^iy. According to the definition of $\tilde{R}^i$, we then have $x\tilde{P}^iy$. Consequently, when individual i's preferences are $\tilde{R}^i$ he strictly prefers x to y and so, given that the others report $\mathbf{R}^{-i}$, individual i strictly prefers lying (reporting R^i and obtaining x) to telling the truth (reporting $\tilde{R}^i$ and obtaining y), contradicting strategy-proofness. We conclude that $c(\tilde{R}^i, \mathbf{R}^{-i}) = x$.

[12] Muller and Satterthwaite (1977) show that strategy-proofness is equivalent to what they call *strong-positive association*, which is equivalent to monotonicity when individual preferences do not display indifference.

Let $(R^1, \ldots, R^N)$ and $(\tilde{R}^1, \ldots, \tilde{R}^N)$ be preference profiles such that $c(R^1, \ldots, R^N) = x$, and such that for every individual i and every $y \in X$ distinct from x, $xR^iy \Longrightarrow x\tilde{P}^iy$. To prove that $c(\cdot)$ is monotonic, we must show that $c(\tilde{R}^1, \ldots, \tilde{R}^N) = x$. But this follows immediately from the result just proven – simply change the preference profile from $(R^1, \ldots, R^N)$ to $(\tilde{R}^1, \ldots, \tilde{R}^N)$ by switching, one at a time, the preferences of each individual i from R^i to $\tilde{R}^i$. We conclude that $c(\cdot)$ is monotonic.

(b) *Pareto Efficiency.* Let x be an arbitrary social state and let $\hat{\mathbf{R}}$ be a preference profile with x at the top of each individual's ranking. We must show that $c(\hat{\mathbf{R}}) = x$.

Because the range of $c(\cdot)$ is all of X, there is some preference profile $\mathbf{R}$ such that $c(\mathbf{R}) = x$. Obtain the preference profile $\tilde{\mathbf{R}}$ from $\mathbf{R}$ by moving x to the top of every individual's ranking. By monotonicity (proven above in (a)), $c(\tilde{\mathbf{R}}) = x$. Because $\hat{\mathbf{R}}$ places x at the top of every individual ranking and $c(\tilde{\mathbf{R}}) = x$, we can again apply monotonicity (do you see why?) and conclude that $c(\hat{\mathbf{R}}) = x$, as desired.

Part 2. $\#X \geq 3 +$ *monotonicity* $+$ *Pareto efficiency* $\Longrightarrow$ *dictatorship.*

The second part of the proof, like our first proof of Arrow's theorem, will use a series of well-chosen preference profiles to uncover a dictator. Given the results from Part 1, we can and will freely use the fact that $c(\cdot)$ is both monotonic and Pareto efficient. Also, in each of the particular figures employed in this proof, all individual rankings are strict. That is, no individual is indifferent between any two social states. We emphasise that this is not an additional assumption – we are *not* ruling out indifference. It just so happens that we are able to provide a proof of the desired result by considering a particular subset of preferences that do not exhibit indifference.

Step 1. Consider any two distinct social states $x, y \in X$ and a profile of strict rankings in which x is ranked highest and y lowest for every individual $i = 1, \ldots, N$. Pareto efficiency implies that the social choice at this profile is x. Consider now changing individual 1's ranking by strictly raising y in it one position at a time. By monotonicity, the social choice remains equal to x so long as y is below x in 1's ranking. But when y finally does rise above x, monotonicity implies that the social choice either changes to y or remains equal to x (see Exercise 6.18(a)). If the latter occurs, then begin the same process with individual 2, then 3, etc. until for some individual n, the social choice does change from x to y when y rises above x in n's ranking. (There must be such an individual n because y will eventually be at the top of every individual's ranking and by Pareto efficiency the social choice will then be y.) Figs. 6.11 and 6.12 depict the situations just before and just after individual n's ranking of y is raised above x.

Step 2. This is perhaps the trickiest step in the proof, so follow closely. Consider Figs. 6.13 and 6.14 below. Fig. 6.13 is derived from Fig. 6.11 (and Fig. 6.14 from Fig. 6.12) by moving x to the bottom of individual i's ranking for $i < n$ and moving it to the second last position in i's ranking for $i > n$. We wish to argue that these changes do not affect the social choices, i.e., that the social choices are as indicated in the figures.

R^1	$\cdots$	R^{n-1}	R^n	R^{n+1}	$\cdots$	R^N	Social Choice
y	$\cdots$	y	x	x	$\cdots$	x	
x	$\cdots$	x	y	$\cdot$		$\cdot$	
$\cdot$		$\cdot$	$\cdot$	$\cdot$		$\cdot$	x
$\cdot$		$\cdot$	$\cdot$	$\cdot$		$\cdot$	
$\cdot$		$\cdot$	$\cdot$	$\cdot$		$\cdot$	
$\cdot$		$\cdot$	$\cdot$	y	$\cdots$	y	

Figure 6.11.

R^1	$\cdots$	R^{n-1}	R^n	R^{n+1}	$\cdots$	R^N	Social Choice
y	$\cdots$	y	y	x	$\cdots$	x	
x	$\cdots$	x	x	$\cdot$		$\cdot$	y
$\cdot$		$\cdot$	$\cdot$	$\cdot$		$\cdot$	
$\cdot$		$\cdot$	$\cdot$	$\cdot$		$\cdot$	
$\cdot$		$\cdot$	$\cdot$	y	$\cdots$	y	

Figure 6.12.

R^1	$\cdots$	R^{n-1}	R^n	R^{n+1}	$\cdots$	R^N	Social Choice
y	$\cdots$	y	x	$\cdot$		$\cdot$	
$\cdot$		$\cdot$	y	$\cdot$		$\cdot$	x
$\cdot$		$\cdot$	$\cdot$	$\cdot$		$\cdot$	
$\cdot$		$\cdot$	$\cdot$	x	$\cdots$	x	
x	$\cdots$	x	$\cdot$	y	$\cdots$	y	

Figure 6.13.

R^1	$\cdots$	R^{n-1}	R^n	R^{n+1}	$\cdots$	R^N	Social Choice
y	$\cdots$	y	y	$\cdot$		$\cdot$	
$\cdot$		$\cdot$	x	$\cdot$		$\cdot$	
$\cdot$		$\cdot$	$\cdot$	$\cdot$		$\cdot$	y
$\cdot$		$\cdot$	$\cdot$	$\cdot$		$\cdot$	
$\cdot$		$\cdot$	$\cdot$	x	$\cdots$	x	
x	$\cdots$	x	$\cdot$	y	$\cdots$	y	

Figure 6.14.

First, note that the social choice in Fig. 6.14 must, by monotonicity, be y because the social choice in Fig. 6.12 is y and no individual's ranking of y versus any other social state changes in the move from Fig. 6.12 to Fig. 6.14 (see Exercise 6.18(b)). Next, note that the profiles in Figs. 6.13 and 6.14 differ only in individual n's ranking of x and y. So, because the social choice in Fig. 6.14 is y, the social choice in Fig. 6.13 must, by monotonicity, be either x or y (we used this same logic in Step 1 – see Exercise 6.18(a)). But

R^1	$\cdots$	R^{n-1}	R^n	R^{n+1}	$\cdots$	R^N	Social Choice
·		·	x	·		·	
·		·	z	·		·	
·		·	y	·		·	x
z	$\cdots$	z	·	z	$\cdots$	z	
y	$\cdots$	y	·	x	$\cdots$	x	
x	$\cdots$	x	·	y	$\cdots$	y	

Figure 6.15.

R^1	$\cdots$	R^{n-1}	R^n	R^{n+1}	$\cdots$	R^N	Social Choice
·		·	x	·		·	
·		·	z	·		·	
·		·	y	·		·	
·		·	·	·		·	x
·		·	·	·		·	
·		·	·	·		·	
z	$\cdots$	z	·	z	$\cdots$	z	
y	$\cdots$	y	·	y	$\cdots$	y	
x	$\cdots$	x	·	x	$\cdots$	x	

Figure 6.16.

if the social choice in Fig. 6.13 is y, then by monotonicity (see Exercise 6.18(b)), the social choice in Fig. 6.11 must be y, a contradiction. Hence, the social choice in Fig. 6.13 is x.

Step 3. Because there are at least three social states, we may consider a social state $z \in X$ distinct from x and y. Since the (otherwise arbitrary) profile of strict rankings in Fig. 6.15 can be obtained from the Fig. 6.13 profile without changing the ranking of x versus any other social state in any individual's ranking, the social choice in Fig. 6.15 must, by monotonicity, be x (see Exercise 6.18(b)).

Step 4. Consider the profile of rankings in Fig. 6.16 derived from the Fig. 6.15 profile by interchanging the ranking of x and y for individuals $i > n$. Because this is the only difference between the profiles in Figs. 6.15 and 6.16, and because the social choice in Fig. 6.15 is x, the social choice in Fig. 6.16 must, by monotonicity, be either x or y (see Exercise 6.18(a)). But the social choice in Fig. 6.16 cannot be y because z is ranked above y in every individual's Fig. 6.16 ranking, and monotonicity would then imply that the social choice would remain y even if z were raised to the top of every individual's ranking, contradicting Pareto efficiency. Hence the social choice in Fig. 6.16 is x.

Step 5. Note that an arbitrary profile of strict rankings with x at the top of individual n's ranking can be obtained from the profile in Fig. 6.16 without reducing the ranking of x versus any other social state in any individual's ranking. Hence, monotonicity (see Exercise 6.18(b)) implies that the social choice must be x whenever individual rankings are strict and x is at the top of individual n's ranking. You are asked to show in Exercise 6.19 that this implies that even when individual rankings are not strict and indifferences are

present, the social choice must be at least as good as x for individual n whenever x is at least as good as every other social state for individual n. So, we may say that individual n is a dictator for the social state x. Because x was arbitrary, we have shown that for each social state $x \in X$, there is a dictator for x. But there cannot be distinct dictators for distinct social states (see Exercise 6.20). Hence there is a single dictator for all social states and therefore the social choice function is dictatorial. ∎

The message you should take away from the Gibbard-Satterthwaite theorem is that, in a rich enough setting, it is impossible to design a non-dictatorial system in which social choices are made based upon self-reported preferences without introducing the possibility that individuals can gain by lying. Fortunately, this does not mean that all is lost. In Chapter 9 we will impose an important and useful domain restriction, known as quasi-linearity, on individual preferences. This will allow us to escape the conclusion of the Gibbard-Satterthwaite theorem and to provide an introduction to aspects of the theory of mechanism design. Thus, the Gibbard-Satterthwaite theorem provides a critically important lesson about the limits of designing systems of social choice based on self-reported information and points us in the direction of what we will find to be rather fertile ground. But before we can develop this further, we must become familiar with the essential and powerful tools of **game theory**, the topic of our next chapter.

6.6 EXERCISES

6.1 Arrow (1951) shows that when the number of alternatives in X is restricted to just *two*, the method of majority voting does yield a social welfare relation that satisfies the conditions of Assumption 6.1. Verify, by example or more general argument, that this is indeed the case.

6.2 Show that the weak Pareto condition *WP* in Arrow's theorem can be replaced with the even weaker Pareto condition *VWP* (very weak Pareto) without affecting the conclusion of Arrow's theorem, where *VWP* is as follows.

VWP. 'If xP^iy for all i, then xPy'.

6.3 (a) Show that the social welfare function that coincides with individual i's preferences satisfies *U*, *WP*, and *IIA*. Call such a social welfare function an *individual i dictatorship*.

(b) Suppose that society ranks any two social states x and y according to individual 1's preferences unless he is indifferent in which case x and y are ranked according to 2's preferences unless he is indifferent, etc. Call the resulting social welfare function a *lexicographic dictatorship*. Show that a lexicographic dictatorship satisfies *U*, *WP* and *IIA* and that it is distinct from an individual i dictatorship.

(c) Describe a social welfare function distinct from an individual i dictatorship and a lexicographic dictatorship that satisfies *U*, *WP* and *IIA*.

6.4 Suppose that X is a non-singleton convex subset of $\mathbb{R}^K$ and that f is a social welfare function satisfying *U* in the sense that it maps every profile of continuous utility functions $\mathbf{u}(\cdot) = (u^1(\cdot), \ldots, u^N(\cdot))$ on X into a continuous social utility function $f_{\mathbf{u}}: X \to \mathbb{R}$. Suppose also that f satisfies *IIA*, *WP*, and *PI*.

Throughout this question you may assume that for any finite number of social states in X and any utility numbers you wish to assign to them, there is a continuous utility function defined on all of X assigning to those states the desired utility numbers. (You might wish to try and prove this. The hints section provides a solution.)

(a) Using *U*, *IIA*, and *PI*, show that if $\mathbf{u}(x) = \mathbf{v}(x')$ and $\mathbf{u}(y) = \mathbf{v}(y')$, then $f_{\mathbf{u}}(x) \geq f_{\mathbf{u}}(y)$ if and only if $f_{\mathbf{v}}(x') \geq f_{\mathbf{v}}(y')$.

Define the binary relation $\succsim$ on $\mathbb{R}^N$ as follows: $(a_1, \ldots, a_N) \succsim (b_1, \ldots, b_N)$ if $f_{\mathbf{u}}(x) \geq f_{\mathbf{u}}(y)$ for some vector of continuous utility functions $\mathbf{u}(\cdot) = (u^1(\cdot), \ldots, u^N(\cdot))$ and some pair of social states x and y satisfying $u^i(x) = a_i$ and $u^i(y) = b_i$ for all i.

(b) Show that $\succsim$ is complete.

(c) Use the fact that f satisfies *WP* to show that $\succsim$ is strictly monotonic.

(d) Use the result from part (a) to show that $\succsim$ is transitive. It is here where at least three social states are needed. (Of course, being non-singleton and convex, X is infinite so that there are many more states than necessary for this step.)

(e) It is possible to prove, using in particular the fact that X is non-singleton and convex, that $\succsim$ is continuous. But the proof is technically demanding. Instead, simply assume that $\succsim$ is continuous and use Theorems 1.1 and 1.3 to prove that there is a continuous and strictly increasing function $W\colon \mathbb{R}^N \to \mathbb{R}$ that represents $\succsim$. (You will need to provide a small argument to adjust for the fact that the domain of W is $\mathbb{R}^N$ while the domain of the utility functions in Chapter 1 is $\mathbb{R}^N_+$.)

(f) Show that for every profile of continuous utility functions $\mathbf{u}(\cdot) = (u^1(\cdot), \ldots, u^N(\cdot))$ on X and all pairs of social states x and y,

$$f_{\mathbf{u}}(x) \geq f_{\mathbf{u}}(y) \text{ if and only if } W(u^1(x), \ldots, u^N(x)) \geq W(u^1(y), \ldots, u^N(y)).$$

6.5 Recall the definition of a lexicographic dictatorship from Exercise 6.3.

(a) Suppose $N = 2$. As in Fig. 6.5, fix a utility vector $(\bar{u}_1, \bar{u}_2)$ in the plane and sketch the sets of utility vectors that are socially preferred, socially worse and socially indifferent to $(\bar{u}_1, \bar{u}_2)$ under a lexicographic dictatorship where individual 1's preferences come first and 2's second. Compare with Fig. 6.5. Pay special attention to the indifference sets.

(b) Conclude from Exercise 6.3 that our first proof of Arrow's theorem does not rule out the possibility of a lexicographic dictatorship and conclude from part (a) of this exercise that our second diagrammatic proof does rule out lexicographic dictatorship. What accounts for the stronger result in the diagrammatic proof?

6.6 In the diagrammatic proof of Arrow's theorem, the claim was made that in Fig. 6.4, we could show either $W(\bar{\mathbf{u}}) < W(\text{IV})$ or $W(\bar{\mathbf{u}}) > W(\text{IV})$. Provide the argument.

6.7 Provide the argument left out of the proof of Theorem 6.2 that the ray starting at $\mathbf{a}$ and extending upward is part of a social indifference curve.

6.8 This exercise considers Theorem 6.2 for the general case of $N \geq 2$. So, let $W\colon \mathbb{R}^N \to \mathbb{R}$ be continuous, strictly increasing and satisfy *HE*.

(a) Suppose that $\min[u^1, \ldots, u^N] = \alpha$. Show that $W(u^1 + \varepsilon, \ldots, u^N + \varepsilon) > W(\alpha, \alpha, \ldots, \alpha)$ for every $\varepsilon > 0$ because W is strictly increasing. Conclude by the continuity of W that $W(u^1, \ldots, u^N) \geq W(\alpha, \alpha, \ldots, \alpha)$.

(b) Suppose that $u^j = \min[u^1, \ldots, u^N] = \alpha$ and that $u^i > \alpha$. Using *HE*, show that $W(\alpha + \varepsilon, u^j, \mathbf{u}^{-ij}) \geq W(u^i, u^j - \varepsilon, \mathbf{u}^{-ij})$ for all $\varepsilon > 0$ sufficiently small, where $\mathbf{u}^{-ij} \in \mathbb{R}^{N-2}$ is the vector $(u^1, \ldots, u^N)$ without coordinates i and j.

(c) Using the continuity of W, conclude from (b) that if $\min[u^1, \ldots, u^N] = \alpha$, then for every individual i, $W(\alpha, \mathbf{u}^{-i}) \geq W(u^1, \ldots, u^N)$, where $\mathbf{u}^{-i} \in \mathbb{R}^{N-1}$ is the vector $(u^1, \ldots, u^N)$ without coordinate i.

(d) By successively applying the result from (c) one individual after another, show that if $\min[u^1, \ldots, u^N] = \alpha$, then $W(\alpha, \alpha, \ldots, \alpha) \geq W(u^1, \ldots, u^N)$.

(e) Using (a) and (d) and the fact that W is strictly increasing, show first that $W(u^1, \ldots, u^N) = W(\tilde{u}^1, \ldots, \tilde{u}^N)$ if and only if $\min(u^1, \ldots, u^N) = \min(\tilde{u}^1, \ldots, \tilde{u}^N)$ and then that $W(u^1, \ldots, u^N) \geq W(\tilde{u}^1, \ldots, \tilde{u}^N)$ if and only if $\min(u^1, \ldots, u^N) \geq \min(\tilde{u}^1, \ldots, \tilde{u}^N)$.

6.9 There are three individuals in society, $\{1, 2, 3\}$, three social states, $\{x, y, z\}$, and the domain of preferences is unrestricted. Suppose that the social preference relation, R, is given by pairwise majority voting (where voters break any indifferences by voting for x first then y then z) if this results in a transitive social order. If this does not result in a transitive social order the social order is $xPyPz$. Let f denote the social welfare function that this defines.

(a) Consider the following profiles, where P^i is individual i's strict preference relation:

Individual 1: xP^1yP^1z

Individual 2: yP^2zP^2x

Individual 3: zP^3xP^3y

What is the social order?

(b) What would be the social order if individual 1's preferences in (a) were instead yP^1zP^1x? or instead zP^1yP^1x?

(c) Prove that f satisfies the Pareto property, *WP*.

(d) Prove that f is non-dictatorial.

(e) Conclude that f does not satisfy *IIA*.

(f) Show directly that f does not satisfy *IIA* by providing two preference profiles and their associated social preferences that are in violation of *IIA*.

6.10 Aggregate income $\bar{y} > 0$ is to be distributed among a set $\mathcal{I}$ of individuals to maximise the utilitarian social welfare function, $W = \sum_{i \in \mathcal{I}} u^i$. Suppose that $u^i = \alpha^i (y^i)^\beta$, where $\alpha^i > 0$ for all $i \in \mathcal{I}$.

(a) Show that if $0 < \beta < 1$, income must be distributed equally if and only if $\alpha^i = \alpha^j$ for all i and j.

(b) Now suppose that $\alpha^i \neq \alpha^j$ for all i and j. What happens in the limit as $\beta \to 0$? How about as $\beta \to 1$? Interpret.

6.11 Suppose utility functions are strictly concave, strictly increasing, and differentiable for every agent in an n-good exchange economy with aggregate endowment $\mathbf{e} \gg \mathbf{0}$.

(a) Show that if $\mathbf{x}^* \gg \mathbf{0}$ is a WEA, then for some suitably chosen weights $\alpha^1, \ldots, \alpha^I > 0$, $\mathbf{x}^*$ maximises the (generalised) utilitarian social welfare function

$$W = \sum_{i \in \mathcal{I}} \alpha^i u^i(\mathbf{x}^i)$$

subject to the resource constraints

$$\sum_{i \in \mathcal{I}} x_j^i \leq \sum_{i \in \mathcal{I}} e_j^i \qquad \text{for } j = i, \ldots, n.$$

(b) Use your findings in part (a) to give an alternative proof of the First Welfare Theorem 5.7.

6.12 The **Borda rule** is commonly used for making collective choices. Let there be N individuals and suppose X contains a finite number of alternatives. Individual i assigns a **Borda count**, $B^i(x)$, to every alternative x, where $B^i(x)$ is the number of alternatives in X to which x is preferred by agent i. Alternatives are then ranked according to their total Borda count as follows:

$$xRy \iff \sum_{i=1}^{N} B^i(x) \geq \sum_{i=1}^{N} B^i(y).$$

(a) Show that the Borda rule satisfies *U*, *WP*, and *D* in Assumption 6.1.

(b) Show that it does *not* satisfy *IIA*.

6.13 Individual i is said to be *decisive* in the social choice between x and y if xP^iy implies xPy, regardless of others' preferences. Sen (1970b) interprets 'liberal values' to imply that there are certain social choices over which each individual should be decisive. For example, in the social choice between individual i's reading or not reading a certain book, the preference of individual i should determine the social preference. Thus, we can view liberalism as a condition on the social welfare relation requiring that every individual be decisive over at least one pair of alternatives. Sen weakens this requirement further, defining a condition he calls **minimal liberalism** as follows:

L^*: there are at least two people k and j and two pairs of distinct alternatives (x, y) and (z, w) such that k and j are decisive over (x, y) and (z, w), respectively.

Prove that there exists *no* social welfare relation that satisfies (merely) the conditions *U*, *WP*, and L^*.

6.14 Atkinson (1970) proposes an index of equality in the distribution of income based on the notion of 'equally distributed equivalent income', denoted y_e. For any strictly increasing, symmetric, and quasiconcave social welfare function over income vectors, $W(y^1, \ldots, y^N)$, income y_e is defined as that amount of income which, if distributed to each individual, would produce the same level of social welfare as the given distribution. Thus, letting $\mathbf{e} \equiv (1, \ldots, 1)$ and $\mathbf{y} \equiv (y^1, \ldots, y^N)$, we have

$$W(y_e\mathbf{e}) \equiv W(\mathbf{y}).$$

Letting μ be the mean of the income distribution $\mathbf{y}$, an index of *equality* in the distribution of income then can be defined as follows:

$$I(\mathbf{y}) \equiv \frac{y_e}{\mu}.$$

(a) Show that $0 < I(\mathbf{y}) \leq 1$ whenever $y_i > 0$ for all i.

(b) Show that the index $I(\mathbf{y})$ is always 'normatively significant' in the sense that for any two income distributions, $\mathbf{y}^1, \mathbf{y}^2$ with the same mean, $I(\mathbf{y}^1)$ is greater than, equal to, or less than $I(\mathbf{y}^2)$ if and only if $W(\mathbf{y}^1)$ is greater than, equal to, or less than $W(\mathbf{y}^2)$, respectively.

6.15 Blackorby and Donaldson (1978) built upon the work of Atkinson described in the preceding exercise. Let $W(\mathbf{y})$ be any strictly increasing, symmetric, and quasiconcave social welfare function defined over income distributions. The authors define a 'homogeneous implicit representation of W' as follows:

$$F(w, \mathbf{y}) \equiv \max_{\lambda}\{\lambda > 0 \mid W(\mathbf{y}/\lambda) \geq w\},$$

where $w \in \mathbb{R}$ is any 'reference level' of the underlying social welfare function. They then define their index of equality in the distribution of income as follows:

$$E(w, \mathbf{y}) \equiv \frac{F(w, \mathbf{y})}{F(w, \mu\mathbf{e})},$$

where, again, μ is the mean of the distribution $\mathbf{y}$ and $\mathbf{e}$ is a vector of 1's.

(a) Show that $F(w, \mathbf{y})$ is homogeneous of degree 1 in the income vector. Show that $F(w, \mathbf{y})$ is greater than, equal to, or less than unity as $W(y)$ is greater than, equal to, or less than w, respectively.

(b) Show that if $W(\mathbf{y})$ is homothetic, $E(w, \mathbf{y})$ is 'reference-level-free' so that $E(w, \mathbf{y}) = E^*(\mathbf{y})$ for all $\mathbf{y}$.

(c) Show that if $W(\mathbf{y})$ is homothetic, $E(w, \mathbf{y}) = I(\mathbf{y})$, where $I(\mathbf{y})$ is the Atkinson index defined in the preceding exercise. Conclude, therefore, that under these conditions, $E(w, \mathbf{y})$ is also normatively significant and lies between zero and 1.

(d) Suppose the social welfare function is the utilitarian form, $W = \sum_{i=1}^{N} y^i$. Show that $E(w, \mathbf{y}) = 1$, denoting 'perfect equality', regardless of the distribution of income. What do you conclude from this?

(e) Derive the index $E(w, \mathbf{y})$ when the social welfare function is the CES form

$$W(\mathbf{y}) = \left(\sum_{i=1}^{N} (y^i)^{\rho}\right)^{1/\rho}, \qquad 0 \neq \rho < 1.$$

6.16 Let $\mathbf{x} \equiv (\mathbf{x}^1, \ldots, \mathbf{x}^N)$ be an allocation of goods to agents, and let the economy's feasible set of allocations be T. Suppose $\mathbf{x}^*$ maximises the utilitarian social welfare function, $W = \sum_{i=1}^{N} u^i(\mathbf{x}^i)$, subject to $\mathbf{x} \in T$.

(a) Let ψ^i for $i = 1, \ldots, N$ be an arbitrary set of increasing functions of one variable. Does $\mathbf{x}^*$ maximise $\sum_{i=1}^{N} \psi^i(u^i(\mathbf{x}^i))$ over $\mathbf{x} \in T$? Why or why not?

(b) If in part (a), $\psi^i = \psi$ for all i, what would your answer be?

(c) If $\psi^i \equiv a^i + b^i u^i(\mathbf{x}^i)$ for arbitrary a^i and $b^i > 0$, what would your answer be?

(d) If $\psi^i \equiv a^i + b u^i(\mathbf{x}^i)$ for arbitrary a^i and $b > 0$, what would your answer be?

(e) How do you account for any similarities and differences in your answers to parts (a) through (d)?

6.17 From the preceding exercise, let $\mathbf{x}^*$ maximise the Rawlsian social welfare function, $W = \min[u^1(\mathbf{x}^1), \ldots, u^N(\mathbf{x}^N)]$ over $\mathbf{x} \in T$.

(a) If ψ^i for $i = 1, \ldots, N$ is an arbitrary set of increasing functions of one variable, must $\mathbf{x}^*$ maximise the function, $\min[\psi^1(u^1(\mathbf{x}^1)), \ldots, \psi^N(u^N(\mathbf{x}^N))]$, over $\mathbf{x} \in T$? Why or why not?

(b) If in part (a), $\psi^i = \psi$ for all i, what would your answer be?

(c) How do you account for your answers to parts (a) and (b)?

(d) How do you account for any differences or similarities in your answers to this exercise and the preceding one?

6.18 Suppose that $c(\cdot)$ is a monotonic social choice function and that $c(\mathbf{R}) = x$, where $R^1, \ldots, R^N$ are each strict rankings of the social states in X.

(a) Suppose that for some individual i, R^i ranks y just below x, and let $\tilde{R}^i$ be identical to R^i except that y is ranked just above x – i.e., the ranking of x and y is reversed. Prove that either $c(\tilde{R}^i, \mathbf{R}^{-i}) = x$ or $c(\tilde{R}^i, \mathbf{R}^{-i}) = y$.

(b) Suppose that $\tilde{R}^1, \ldots, \tilde{R}^N$ are strict rankings such that for every individual i, the ranking of x versus any other social state is the same under $\tilde{R}^i$ as it is under R^i. Prove that $c(\tilde{\mathbf{R}}) = x$.

6.19 Let $c(\cdot)$ be a monotonic social choice function and suppose that the social choice must be x whenever all individual rankings are strict and x is at the top of individual n's ranking. Show the social choice must be at least as good as x for individual n when the individual rankings are not necessarily strict and x is at least as good for individual n as any other social state.

6.20 Let x and y be distinct social states. Suppose that the social choice is at least as good as x for individual i whenever x is at least as good as every other social state for i. Suppose also that the social choice is at least as good as y for individual j whenever y is at least as good as every other social state for j. Prove that $i = j$.

6.21 Call a social choice function *strongly monotonic* if $c(\mathbf{R}) = x$ implies $c(\tilde{\mathbf{R}}) = x$ whenever for every individual i and every $y \in X$, $xR^iy \Longrightarrow x\tilde{R}^iy$.

Suppose there are two individuals, 1 and 2, and three social states, x, y, and z. Define the social choice function $c(\cdot)$ to choose individual 1's top-ranked social state unless it is not unique, in which case the social choice is individual 2's top-ranked social state among those that are top-ranked for individual 1, unless this too is not unique, in which case, among those that are top-ranked for both individuals, choose x if it is among them, otherwise choose y.

(a) Prove that $c(\cdot)$ is strategy-proof.

(b) Show by example that $c(\cdot)$ is not strongly monotonic. (Hence, strategy-proofness does not imply strong monotonicity, even though it implies monotonicity.)

6.22 Show that if $c(\cdot)$ is a monotonic social choice function and the finite set of social states is X, then for every $x \in X$ there is a profile, $\mathbf{R}$, of *strict* rankings such that $c(\mathbf{R}) = x$. (Recall that, by definition, every x in X is chosen by $c(\cdot)$ at some preference profile.)

6.23 Show that when there are just two alternatives and an odd number of individuals, the majority rule social choice function (i.e., that which chooses the outcome that is the top ranked choice for the majority of individuals) is Pareto efficient, strategy-proof and non-dictatorial.