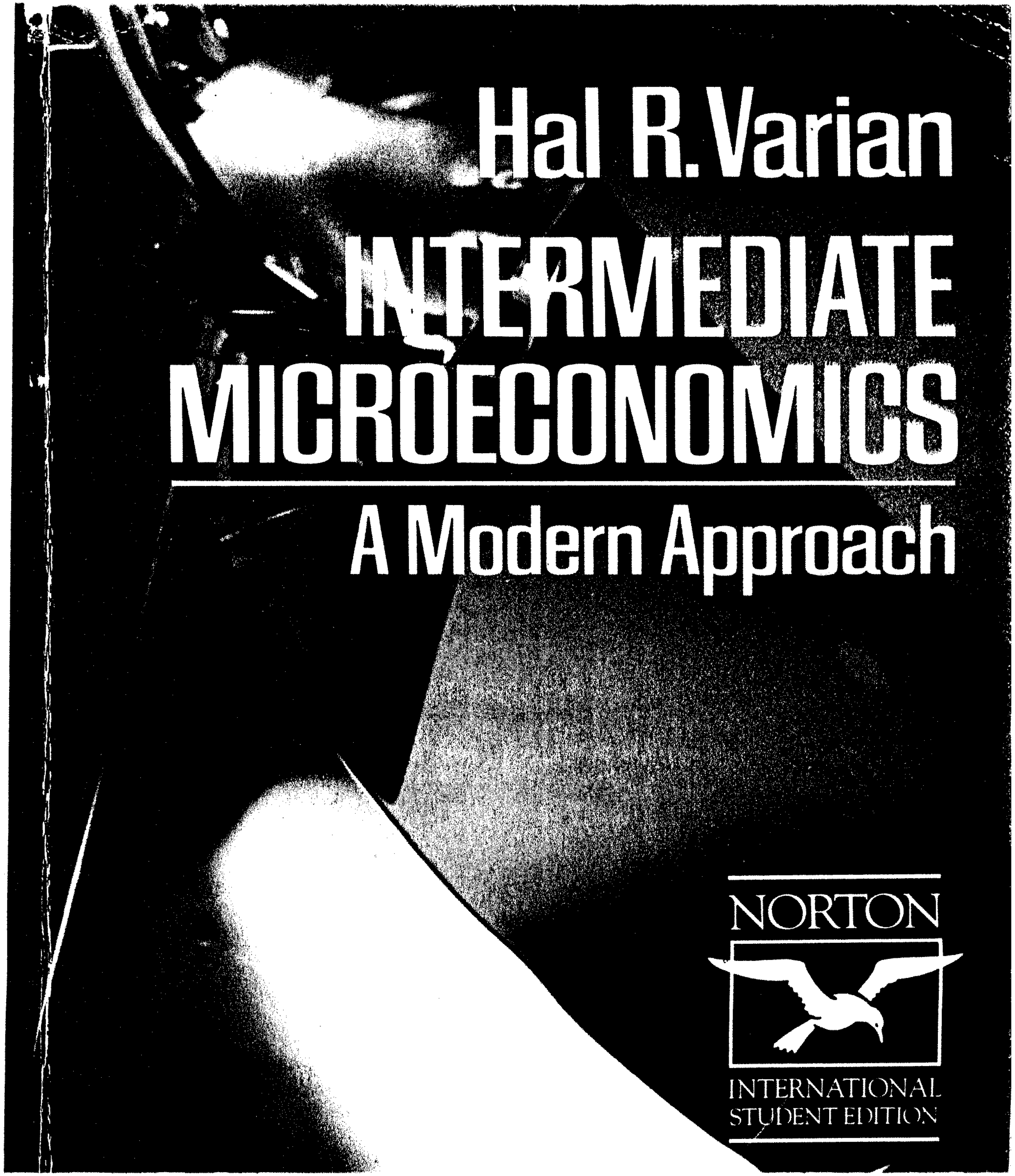
Hal R. Varian

# INTERMEDIATE MICROECONOMICS

## A Modern Approach



NORTON

INTERNATIONAL STUDENT EDITION

# 30

# WELFARE

Up until now we have focused on considerations of Pareto efficiency in evaluating economic allocations. But there are other important considerations. It must be remembered that Pareto efficiency has nothing to say about the distribution of welfare across people; giving everything to one person will typically be Pareto efficient. But the rest of us might not consider this a reasonable allocation. In this chapter we will investigate some techniques that can be used to formalize ideas related to the distribution of welfare.

Pareto efficiency is in itself a desirable goal—if there is some way to make some group of people better off without hurting other people, why not do it? But there will usually be many Pareto efficient allocations; how can society choose among them?

The major focus of this chapter will be the idea of a **welfare function**, which provides a way to "add together" different consumer's utilities. More generally, a welfare function provides a way to rank different distributions of utility among consumers. Before we investigate the implications of this concept, it is worthwhile considering just how one might go about "adding together" the individual consumers' preferences to construct some kind of "social preferences."

## 30.1 Aggregation of Preferences

Let us return to our early discussion of consumer preferences. As usual, we will assume that these preferences are transitive. Originally, we thought of a consumer's preferences as being defined over his own bundle of goods, but now we want to expand on that concept and think of each consumer having preferences over the entire allocation of goods among the consumers. Of course, this includes the possibility that the consumer might not care about what other people have, just as we had originally assumed.

Let us use the symbol **x** to denote a particular allocation—a description of what every individual gets of every good. Then given two allocations, **x** and **y**, each individual $i$ can say whether or not he or she prefers **x** to **y**.

Given the preferences of all the agents, we would like to have a way to "aggregate" them into one **social preference.** That is, if we know how all the individuals rank various allocations, we would like to be able to use this information to develop a social ranking of the various allocations. This is the problem of social decision making at its most general level. Let's consider a few examples.

One way to aggregate individual preferences is to use some kind of voting. We could agree that **x** is "socially preferred" to **y** if a majority of the individuals prefer **x** to **y**. However, there is a problem with this method—it may not generate a transitive social preference ordering. Consider, for example, the case illustrated in Table 30.1.

**Table 30.1** **Preferences that lead to intransitive voting.**

| Person A | Person B | Person C |
|---|---|---|
| x | y | z |
| y | z | x |
| z | x | y |

Here we have listed the rankings for three alternatives, **x**, **y**, and **z**, by three people. Note that a majority of the people prefer **x** to **y**, a majority prefer **y** to **z**, and a majority prefer **z** to **x**. Thus aggregating individual preferences by majority vote won't work since, in general, the social preferences resulting from majority voting aren't well-behaved preferences, since they are not transitive. Since the preferences aren't transitive, there will be no "best" alternative from the set of alternatives (**x**, **y**, **z**). Which outcome society chooses will depend on the order in which the vote is taken.



To see this suppose that the three people depicted in Table 30.1 decide to vote first on **x** versus **y**, and then vote on the winner of this contest versus **z**. Since a majority prefer **x** to **y**, the second contest will be between **x** and **z**, which means that **z** will be the outcome.

But what if they decide to vote on **z** versus **x** and then pit the winner of this vote against **y**? Now **z** wins the first vote, but **y** beats **z** in the second vote. Which outcome is the overall winner depends crucially on the order in which the alternatives are presented to the voters.

Another kind of voting mechanism that we might consider is rank-order voting. Here each person ranks the goods according to his preferences and assigns a number that indicates its rank in his ordering: for example, a 1 for the best alternative, 2 for the second best, and so on. Then we sum up the scores of each alternative across the people to determine an aggregate score for each alternative, and say that one outcome is socially preferred to another if it has a lower score.

In Table 30.2 we have illustrated a possible preference ordering for three allocations **x**, **y**, and **z** by two people. Suppose first that only alternatives **x** and **y** were available. Then in this example **x** would be given a rank of 1 by person A and 2 by person B. The alternative **y** would be given just the reverse ranking. Thus the outcome of the voting would be a tie with each alternative having an aggregate rank of 3.

**Table 30.2** **The choice between x and y depends on z.**

| Person A | Person B |
|---|---|
| x | y |
| y | z |
| z | x |

But now suppose that **z** is introduced to the ballot. Person A would give **x** a score of 1, **y** a score of 2, and **z** a rank of 3. Person B would give **y** a score of 1, **z** a score of 2, and **x** a score of 3. This means that **x** would now have an aggregate rank of 4, and **y** would have an aggregate rank of 3. In this case **y** would be preferred to **x** by rank-order voting.

The problem with both majority voting and rank-order voting is that their outcomes can be manipulated by astute agents. Majority voting can be manipulated by changing the order on which things are voted so as to yield the desired outcome. Rank-order voting can be manipulated by introducing new alternatives that change the final ranks of the relevant alternatives.

The question naturally arises as to whether there are social decision mechanisms—ways of aggregating preferences—that are immune to this kind of manipulation? Are there ways to "add up" preferences that don't have the undesirable properties described above?

Let's list some things that we would want our social decision mechanism to do:

1. Given any set of complete, reflexive, and transitive individual preferences, the social decision mechanism should result in social preferences that satisfy the same properties.

2. If everybody prefers alternative **x** to alternative **y**, then the social preferences should rank **x** ahead of **y**.

3. The preferences between **x** and **y** should depend only on how people rank **x** versus **y**, and not on how they rank other alternatives.

All three of these requirements seem eminently plausible. Yet it can be quite difficult to find a mechanism that satisfies all of them. In fact, Kenneth Arrow has proved the following remarkable result:[1]

**Arrow's Impossibility Theorem.** *If a social decision mechanism satisfies properties 1, 2, and 3, then it must be a dictatorship: all social rankings are the rankings of one individual.*

Arrow's Impossibility Theorem is quite surprising. It shows that three very plausible and desirable features of a social decision mechanism are inconsistent with democracy: there is no "perfect" way to make social decisions. There is no perfect way to "aggregate" individual preferences to make one social preference. If we want to find a way to aggregate individual preferences to form social preferences, we will have to give up one of the properties of a social decision mechanism described in Arrow's theorem.

## 30.2 Social Welfare Functions

If we were to drop any of the desired features of a social welfare function described above, it would probably be property 3—that the social preferences between two alternatives only depends on the ranking of those two alternatives. If we do that, certain kinds of rank-order voting become possibilities.

[1] See Kenneth Arrow, *Social Choice and Individual Values* (New York: Wiley, 1963). Arrow, a professor at Stanford University, was awarded the Nobel Prize in economics for his work in this area.



Given the preferences of each individual $i$ over the allocations, we can construct utility functions, $u_i(\mathbf{x})$, that summarize the individuals' value judgments: person $i$ prefers $\mathbf{x}$ to $\mathbf{y}$ if and only if $u_i(\mathbf{x}) > u_i(\mathbf{y})$. Of course, these are just like all utility functions—they can be scaled in any way that preserves the underlying preference ordering. There is no *unique* utility representation.

But let us pick some utility representation and stick with it. Then one way of getting social preferences from individuals' preferences is to add up the individual utilities and use the resulting number as a kind of social utility. That is, we will say that allocation **x** is socially preferred to allocation **y** if

$$\sum_{i=1}^{n} u_i(\mathbf{x}) > \sum_{i=1}^{n} u_i(\mathbf{y})$$

where $n$ is the number of individuals in the society.

This works—but of course it is totally arbitrary, since our choice of utility representation is totally arbitrary. The choice of using the sum is also arbitrary. Why not use a weighted sum of utilities? Why not use the product of utilities, or the sum of the squares of utilities?

One reasonable restriction that we might place on the "aggregating function" is that it be increasing in each individual's utility. That way we are assured that if everybody prefers **x** to **y**, then the social preferences will prefer **x** to **y**.

There is a name for this kind of aggregating function; it is called a **social welfare function.** A social welfare function is just some function of the individual utility functions: $W(u_1(\mathbf{x}), \ldots, u_n(\mathbf{x}))$. It gives a way to rank different allocations that depends only on the individual preferences, and it is an increasing function of each individual's utility.

Let's look at some examples. One special case mentioned above is the *sum* of the individual utility functions

$$W(u_1, \ldots, u_n) = \sum_{i=1}^{n} u_i.$$

This is sometimes referred to as a **classical utilitarian** or **Benthamite** welfare function. [2] A slight generalization of this form is the **weighted-sum-of-utilities** welfare function:

$$W(u_1, \ldots, u_n) = \sum_{i=1}^{n} a_i u_i.$$

[2] Jeremy Bentham (1748–1832) was the founder of the utilitarian school of moral philosophy, a school that considers the highest good to be the greatest happiness for the greatest number.

Here the weights, $a_1, \ldots, a_n$, are supposed to be numbers indicating how important each agent's utility is to the overall social welfare. It is natural to take each $a_i$ as being positive.

Another interesting welfare function is the **minimax** or **Rawlsian** social welfare function:

$$W(u_1, \ldots, u_n) = \min\{u_1, \ldots, u_n\}.$$

This welfare function says that the social welfare of an allocation depends only on the welfare of the worst off agent—the person with the minimal utility.[3]

Each of these is a possible way to compare individual utility functions. Each of them represents different ethical judgments about the comparison between different agents' welfares. About the only restriction that we will place on the structure of the welfare function at this point is that it be increasing in each consumer's utility.

## 30.3 Welfare Maximization

Once we have a welfare function we can examine the problem of welfare maximization. Let us use the notation $x_i^j$ to indicate how much individual $i$ has of good $j$, and suppose that there are $n$ consumers and $k$ goods. Then the allocation $\mathbf{x}$ consists of the list of how much each of the agents has of each of the goods.

If we have a total amount $X^1, \ldots, X^k$ of goods $1, \ldots, k$ to distribute among the consumers, we can pose the welfare maximization problem:

$$\max\ W(u_1(\mathbf{x}), \ldots, u_n(\mathbf{x}))$$

$$\text{such that } \sum_{i=1}^{n} x_i^1 = X^1$$

$$\vdots$$

$$\sum_{i=1}^{n} x_i^k = X^k.$$

Thus we are trying to find the feasible allocation that maximizes social welfare. What properties does such an allocation have?

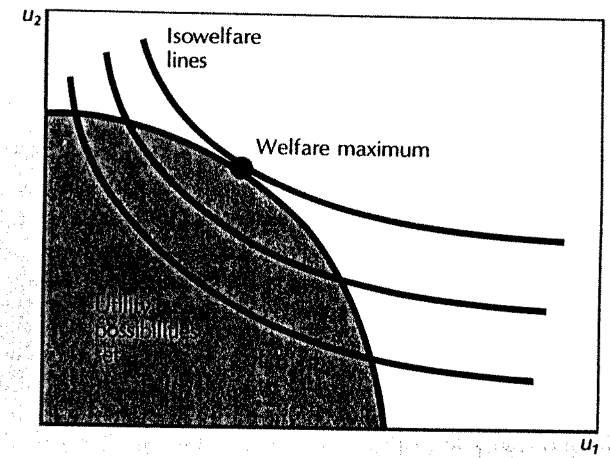
The first thing that we should note is that a maximal welfare allocation must be a Pareto efficient allocation. The proof is easy: suppose that

[3] John Rawls is a contemporary moral philosopher at Harvard who has argued for this principle of justice.



it were not. Then there would be some other feasible allocation that gave everyone at least as large a utility, and someone strictly greater utility. But the welfare function is an increasing function of each agent's utility. Thus this new allocation would have to have higher welfare, which contradicts the assumption that we originally had a welfare maximum.

We can illustrate this situation in Figure 30.1, where the set $U$ indicates the set of possible utilities in the case of two individuals. This set is known as the **utility possibilities set**. The boundary of this set—the **utility possibilities frontier**—is the set of utility levels associated with Pareto efficient allocations. If an allocation is on the boundary of the utility possibilities set, then there are no other feasible allocations that yield higher utilities for both agents.




**Welfare maximization.** An allocation that maximizes a welfare function must be Pareto efficient. **Figure 30.1**

The "indifference curves" in this diagram are called **isowelfare** lines since they depict those distributions of utility that have constant welfare. As usual, the optimal point is characterized by a tangency condition. But for our purposes, the notable thing about this maximal welfare point is that it is Pareto efficient—it must occur on the boundary of the utility possibilities set.

The next observation we can make from this diagram is that *any* Pareto efficient allocation must be a welfare maximum for some welfare function. An example is given in Figure 30.2.

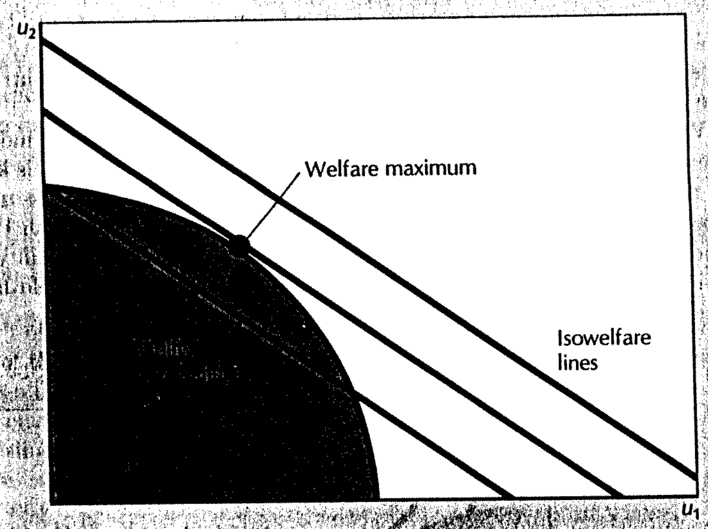



Figure 30.2 **Maximization of the weighted-sum-of-utilities welfare function.** If the utility possibility set is convex, then every Pareto efficient point is a maximum for a weighted-sum-of-utilities welfare function.

In Figure 30.2 we have picked a Pareto efficient allocation and found a set of isowelfare curves for which it is maximal. Actually, we can say a bit more than this. If the set of possible utility distributions is a convex set, as illustrated, then every point on its frontier is a utility maximum for a weighted-sum-of-utilities welfare function, as illustrated in Figure 30.2. The welfare function thus provides a way to single out Pareto efficient allocations: every welfare maximum is a Pareto efficient allocation, and every Pareto efficient allocation is a welfare maximum.

## 30.4 Individualistic Social Welfare Functions

Up until now we have been thinking of individual preferences as being defined over entire allocations rather than over each individual's bundle of goods. But, as we remarked earlier, individuals might only care about their own bundles. In this case, we can use $x_i$ to denote individual *i's* consumption bundle, and let $u_i(x_i)$ be individual *i's* utility level using some fixed representation of utility. Then a social welfare function will have the form

$$W = W(u_1(x_1), \ldots, u_n(x_n)).$$



The welfare function is directly a function of the distribution of utilities, but it is indirectly a function of the individual agents consumption bundles. This special form of welfare function is known as an **individualistic welfare function** or a **Bergson-Samuelson welfare function.**[4]

If each agent's utility depends only on his or her own consumption, then there are no consumption externalities. Thus the standard results of Chapter 28 apply and we have an intimate relationship between Pareto efficient allocations and market equilibria: all competitive equilibria are Pareto efficient, and, under appropriate convexity assumptions, all Pareto efficient allocations are competitive equilibria.

Now we can carry this categorization one step further. Given the relationship between Pareto efficiency and welfare maxima described above, we can conclude that all welfare maxima are competitive equilibria, and all competitive equilibria are welfare maxima for some welfare function.

## 30.5 Fair Allocations

The welfare function approach is a very general way to describe social welfare. But because it is so general it can be used to summarize the properties of many kinds of moral judgments. On the other hand, it isn't much use in deciding what kinds of ethical judgments might be reasonable ones.

Another approach is to start with some specific moral judgments and then examine their implications for economic distribution. This is the approach taken in the study of **fair allocations.** We start with a definition of what might be considered a fair way to divide a bundle of goods, and then use our understanding of economic analysis to investigate its implications.

Suppose that you were given some goods to divide fairly among $n$ equally deserving people. How would you do it? It is probably safe to say that in this problem most people would divide the goods equally among the $n$ agents. Given that they are by hypothesis equally deserving, what else could you do?

What is appealing about this idea of equal division? One appealing feature is that it is *symmetric.* Each agent has the same bundle of goods; no agent prefers any other agent's bundle of goods to his or her own, since they all have exactly the same thing.

Unfortunately, an equal division will not necessarily be Pareto efficient. If agents have different tastes they will generally desire to trade away from equal division. Let us suppose that this trade takes place, and that it moves us to a Pareto efficient allocation.

[4] Abram Bergson and Paul Samuelson are contemporary economists who investigated properties of this kind of welfare function in the early 1940s. Samuelson was awarded a Nobel Prize in economics for his many contributions.

The question arises: is this Pareto efficient allocation still fair in any sense? Does trade from equal division inherit any of the symmetry of the starting point?

The answer is: not necessarily. Consider the following example. We have three people, A, B, and C. A and B have the same tastes, and C has different tastes. We start from an equal division and suppose that A and C get together and trade. Then they will typically both be made better off. Now B, who didn't have the opportunity to trade with C, will **envy** A—that is, he would prefer A's bundle to his own. Even though A and B started with the same allocation, A was luckier in her trading, and this destroyed the symmetry of the original allocation.

This means that arbitrary trading from an equal division will not necessarily preserve the symmetry of the starting point of equal division. We might well ask if there is any allocation that preserves this symmetry? Is there any way to get an allocation that is both Pareto efficient and equitable at the same time?

## 30.6 Envy and Equity

Let us now try to formalize some of these ideas. What do we mean by "symmetric" or "equitable" anyway? One possible set of definitions is as follows.

We say an allocation is **equitable** if no agent prefers any other agent's bundle of goods to his or her own. If some agent $i$ does prefer some other agent $j$'s bundle of goods, we say that $i$ **envies** $j$. Finally, if an allocation is both equitable and Pareto efficient, we will say that it is a **fair** allocation.

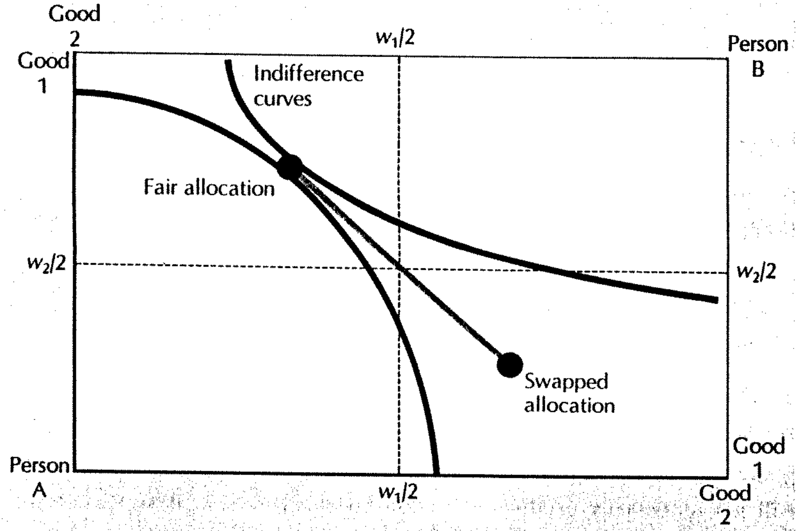
These are ways of formalizing the idea of symmetry alluded to above. An equal division allocation has the property that no agent envies any other agent—but there are many other allocations that have this same property.

Consider Figure 30.3. To determine whether any allocation is equitable or not, just look at the allocation that results if the two agents swap bundles. If this swapped allocation lies "below" each agent's indifference curve through the original allocation, then the original allocation is an equitable allocation. (Here "below" means below from the point of view of each agent; from our point of view the swapped allocation must lie between the two indifference curves.)

Note also that the allocation in Figure 30.3 is also Pareto efficient. Thus it is not only equitable, in the sense that we defined the term, but it is also efficient. By our definition, it is a fair allocation. Is this kind of allocation a fluke, or will fair allocations typically exist?

It turns out that fair allocations *will* generally exist, and there is an easy way to see that this is so. We start with the setup described in the last section, where we have an equal division allocation and consider trading to a Pareto efficient allocation. Instead of using just any old way to trade, let






**Fair allocations.** A fair allocation in an Edgeworth box. Each person prefers the fair allocation to the swapped allocation.

**Figure 30.3**

us use the special mechanism of the competitive market. This will move us to a new allocation where each agent is choosing the best bundle of goods he or she can afford at the equilibrium prices $(p_1, p_2)$, and we know from Chapter 28 that such an allocation must be Pareto efficient.

But is it still equitable? Well, suppose not. Suppose that one of the consumers, say consumer A, envies consumer B. This means that A prefers what B has to her own bundle. In symbols:

$$(x_A^1, x_A^2) \prec_A (x_B^1, x_B^2).$$

But, if A prefers B's bundle to her own, and her own bundle is the best bundle she can afford at the prices $(p_1, p_2)$, this means that B's bundle must cost more than A can afford. In symbols:

$$p_1\omega_A^1 + p_2\omega_A^2 < p_1 x_B^1 + p_2 x_B^2.$$

But this is a contradiction! For by hypothesis, A and B started with exactly the same bundle, since they started from an equal division. If A can't afford B's bundle, then B can't afford it either.

Thus we can conclude that it is impossible for A to envy B in these circumstances. A competitive equilibrium from equal division must be a fair allocation. Thus the market mechanism will preserve certain kinds of equity: if the original allocation is equally divided, the final allocation must be fair.

## Summary

1. Arrow's Impossibility Theorem shows that there is no ideal way to aggregate individual preferences into social preferences.

2. Nevertheless, economists often use welfare functions of one sort or another to represent distributional judgments about allocations.

3. As long as the welfare function is increasing in each individual's utility, a welfare maximum will be Pareto efficient. Furthermore, every Pareto efficient allocation can be thought of as maximizing some welfare function.

4. The idea of fair allocations provides an alternative way to make distributional judgments. This idea emphasizes the idea of symmetric treatment.

5. Even when the initial allocation is symmetric, arbitrary methods of trade will not necessarily produce a fair allocation. However, it turns out that the market mechanism will provide a fair allocation.

## Review Questions

1. Suppose that we say that an allocation **x** is socially preferred to an allocation **y** only if *everyone* prefers **x** to **y**. (This is sometimes called the Pareto ordering, since it is closely related to the idea of Pareto efficiency.) What shortcoming does this have as a rule for making social decisions?

2. The opposite of the Rawlsian welfare function might be called the "Nietzschian" welfare function—a welfare function that says the value of an allocation depends only on the welfare of the *best off* agent. What mathematical form would the Nietzschian welfare function take?

3. Suppose that the utility possibilities set is a convex set, and that consumers care only about their own consumption. What kind of allocations represent welfare maxima of the Nietzschian welfare function?

4. Suppose that an allocation is Pareto efficient, and that each individual only cares about his own consumption. Prove that there must be some individual that envies no one, in the sense described in the text. (This puzzle requires some thought, but it is worth it.)

5. The ability to set the voting agenda can often be a powerful asset. Assuming that social preferences are decided by pair-wise majority voting and that the preferences given in Table 30.1 hold, demonstrate this fact by producing a voting agenda that results in allocation **y** winning. Find an agenda that has **z** as the winner. What property of the social preferences is responsible for this agenda setting power?



## APPENDIX

Here we consider the problem of welfare maximization, using an individualistic welfare function. Using the transformation function described in Chapter 29 to describe the production possibilities frontier, we write the welfare maximization problem as

$$\max_{x_A^1, x_A^2, x_B^1, x_B^2} W(u_A(x_A^1, x_A^2), u_B(x_B^1, x_B^2))$$
$$\text{such that}$$
$$T(X^1, X^2) = 0$$

where we use $X^1$ and $X^2$ to denote the total amount of good 1 and good 2 produced and consumed.

The Lagrangian for this problem is

$$L = W(u_A(x_A^1, x_A^2), u_B(x_B^1, x_B^2)) - \lambda(T(X^1, X^2) - 0).$$

Differentiating with respect to each of the choice variables gives us the first-order conditions

$$\frac{\partial L}{\partial x_A^1} = \frac{\partial W}{\partial u_A}\frac{\partial u_A(x_A^1, x_A^2)}{\partial x_A^1} - \lambda\frac{\partial T(X^1, X^2)}{\partial X^1} = 0$$
$$\frac{\partial L}{\partial x_A^2} = \frac{\partial W}{\partial u_A}\frac{\partial u_A(x_A^1, x_A^2)}{\partial x_A^2} - \lambda\frac{\partial T(X^1, X^2)}{\partial X^2} = 0$$
$$\frac{\partial L}{\partial x_B^1} = \frac{\partial W}{\partial u_B}\frac{\partial u_B(x_B^1, x_B^2)}{\partial x_B^1} - \lambda\frac{\partial T(X^1, X^2)}{\partial X^1} = 0$$
$$\frac{\partial L}{\partial x_B^2} = \frac{\partial W}{\partial u_B}\frac{\partial u_B(x_B^1, x_B^2)}{\partial x_B^2} - \lambda\frac{\partial T(X^1, X^2)}{\partial X^2} = 0.$$

Rearranging and dividing the first equation by the second, and the third by the fourth, we have

$$\frac{\partial u_A/\partial x_A^1}{\partial u_A/\partial x_A^2} = \frac{\partial T/\partial X^1}{\partial T/\partial X^2}$$
$$\frac{\partial u_B/\partial x_B^1}{\partial u_B/\partial x_B^2} = \frac{\partial T/\partial X^1}{\partial T/\partial X^2}.$$

Note that these are exactly the same equations that we encountered in the Appendix to Chapter 29. Thus the welfare maximization problem gives us the same first-order conditions as the Pareto efficiency problem.

This is obviously no accident. According to the discussion in the text, the allocation resulting from the maximization of a Bergson-Samuelson welfare function is Pareto efficient, and every Pareto efficient allocation maximizes some welfare function. Thus welfare maxima and Pareto efficient allocations have to satisfy the same first-order conditions.