2. Pareto optimality and the Pareto criterion

Determination of economic criteria for public policy evaluation has been a subject of great debate. The difficulty stems from the inability to decide on purely economic grounds how the goods and services produced in an economy should be distributed among individuals. Issues of distribution and equity are political and moral as well as economic in nature.

Classical economists such as Bentham (1961, first published 1823) long ago developed the concept of a social welfare function to measure the welfare of society as a function of the utilities of all individuals. The objective was to establish a complete *social ordering* of all possible alternative states of the world. A social ordering, in principle, permits comparison and choice among alternative states and would allow economists to determine precisely which set of policies maximize the good of society. The problem is that agreement on the form of a social welfare function cannot be reached so the use of such a concept has been clouded with controversy. Many functional forms have been proposed and defended on moral, ethical and philosophical grounds with specific considerations given to equity, liberty and justice (see Sections 3.3 and 3.4 for more details). Because of the subjectivity of these arguments, agreement is unlikely ever to be reached. Even if agreement were reached among economists, policy-makers may be unwilling to accept such judgments by economists as the basis for public policy choice.

Because use of a social welfare function is clouded by controversy, many economists have tried to maintain objectivity and the claim of their professional practice as a science by avoiding *value judgments*. A value judgment is simply a subjective statement about what is of value to society that helps to determine the social ordering of alternative states of the world. It is subjective in the sense that it cannot be totally supported by evidence. It is not a judgment of fact. The attempt to avoid value judgments led to development of the *Pareto principle*.

The Pareto criterion was introduced in the nineteenth century by the eminent Italian economist, Vilfredo Pareto (1896). Its potential for application to public policy choices, however, is still very much discussed. By this criterion, a policy change is socially desirable if, by the change, everyone can be made better off, or at least some are made better off, while no one is made worse off. If there are any who lose, the criterion is not met. In his book *The Zero-Sum Society: Distribution and the Possibilities for Economic Change*, Lester Thurow (1980) contends that many good projects do not get under way simply because project managers are unwilling to pay compensation to those who would actually be made worse off. If this is correct, perhaps policy measures should be considered that meet the Pareto criterion. That is, perhaps policy measures that include the payment of compensation, so that everyone is made better off, should be considered. For example, those who support tariffs argue that their removal results in short-term loss of jobs for which workers are not adequately compensated. Trade theory shows that there are economic gains from free trade, but the distribution of these gains is what the workers object to. This objection would probably not arise if only policies that met the Pareto criterion were considered. However, as will become clear, there are also limitations to using the Pareto criterion to rank policy choices.

A large part of theoretical welfare economics and its application is based on the Pareto principle and the concept of Pareto optimality. This chapter discusses both Pareto optimality and the Pareto criterion in a general equilibrium setting. The consideration of these concepts in a general equilibrium context enables greater understanding of the assumptions, limitations and generalizations associated with applying welfare economics to real-world problems discussed in later chapters.

2.1 PARETO OPTIMALITY AND THE PARETO CRITERION DEFINED

The *Pareto criterion* is a technique for comparing or ranking alternative states of the economy. By this criterion, if it is possible to make at least one person better off when moving from state *A* to state *B* without making anyone else worse off, state *B* is ranked higher by society than state *A*. If this is the case, a movement from state *A* to state *B* represents a *Pareto improvement*, or state *B* is *Pareto superior* to state *A*. As an example, suppose a new technology is introduced that causes lower food prices and, at the same time, does not harm anyone by (for example) causing unemployment or reduced profits. The introduction of such a technology would be a Pareto improvement.

To say that society should make movements that are Pareto improvements is, of course, a value judgment but one that enjoys widespread acceptance. Some would disagree, however, if policies continuously make the rich richer while the poor remain unaffected.

If society finds itself in a position from which there is no feasible Pareto improvement, such a state is called a *Pareto optimum*. That is, a *Pareto-optimal state is defined as a state from which it is impossible to make one person better off without making another person worse off*. It is important to stress that, even though a Pareto-optimal state is reached, this in no way implies that society is equitable in terms of income distribution. For example, as will become evident later, a Pareto-optimum position is consistent with a state of nature even where the distribution of income is highly skewed.

If the economy is not at a Pareto optimum, there is some inefficiency in the system. When output is divisible, it is always theoretically possible to make everyone better off in moving from a *Pareto-inferior* position to a *Pareto-superior* position. Hence, Pareto-optimal states are also referred to as *Pareto-efficient* states, and the Pareto criterion is referred to as an *efficiency criterion*. Efficiency in this context is associated with getting as much as possible for society from its limited resources. Note, however, that the Pareto criterion can be used to compare two inefficient states as well. That is, one inefficient state may represent a Pareto improvement over another inefficient state.

Of course, it may be politically infeasible to move from certain inefficient states to certain Pareto-superior states. If Thurow is correct, the only feasible options may be moves to positions where at least one person is made worse off. States where one person is made better off and another is made worse off are referred to as *Pareto-noncomparable states*. Note that all of these Pareto-related concepts are defined independently of societies' institutional arrangements for production, marketing and trade.

2.2 THE PURE CONSUMPTION CASE

Now consider the concepts of Pareto optimality and the Pareto criterion for the pure exchange case, that is, the optimal allocation of goods among individuals where the goods are, in fact, already produced. In this context, a set of marginal exchange conditions characterizing Pareto-efficient states can be developed. Suppose that there are two individuals, A and B, and quantities of two goods, $\bar{q}_1$ and $\bar{q}_2$, which have been produced and can be distributed between the two individuals. This situation is represented by the *Edgeworth–Bowley box* in Figure 2.1, where the width of the box measures the total amount of $\bar{q}_1$ produced and the height of the box measures the total amount of $\bar{q}_2$ produced.

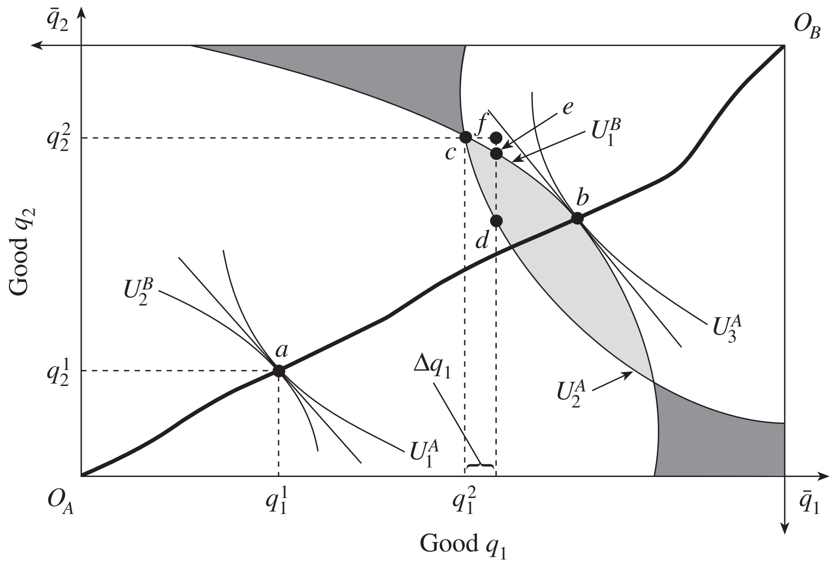

Figure 2.1

The indifference map for individual A is displayed in the box in standard form with O_A as the origin. Three indifference curves for individual A – labeled U_1^A, U_2^A, and U_3^A – are drawn in the box. The indifference map for individual B with indifference curves U_1^B and U_2^B has O_B as the origin and thus appears upside down and reversed. Displaying the indifference maps of individuals A and B in this manner ensures that every point in the box represents a particular distribution of $\bar{q}_1$ and $\bar{q}_2$ or a *state of the economy* in the pure exchange case. For example, at point a, individual A is endowed with q_1^1 of good q_1 and q_2^1 of good q_2, while the remainder of each, $\bar{q}_1 - q_1^1$ and $\bar{q}_2 - q_2^1$, respectively, is distributed to individual B.

The solid line in Figure 2.1 running from O_A to O_B through points a and b is known as the *contract curve* and is constructed by connecting all points of tangency between indifference curves for individuals A and B. At all points on this line, both consumers' indifference curves have the same slope or, in other words, both consumers have equal *marginal rates of substitution* for goods q_1 and q_2. The marginal rate of substitution measures the rate at which a consumer is willing to trade one good for another at the margin.

The marginal rate of substitution for each consumer generally varies along the contract curve for both consumers. The slope of the indifference curves at point *a*, for example, is not necessarily the same as the slope at point *b*.

Now, consider the possibility of using the Pareto criterion to compare or rank alternative states of the economy. For example, compare point *c* with other points in the Edgeworth-Bowley box. First, compare point *c* with points inside the lightly shaded area. At point *c*, the marginal rate of substitution of q_1 for q_2 for individual *A*, denoted by $\mathrm{MRS}^A_{q_1q_2}$ is greater than the marginal rate of substitution of q_1 for q_2 for individual *B*, denoted by $\mathrm{MRS}^B_{q_1q_2}$. This implies that the amount of q_2 that individual *A* is willing to give up to obtain an additional unit of q_1 exceeds the amount individual *B* is willing to accept to give up a unit of q_1. For the marginal increment Δq_1 in Figure 2.1, this excess corresponds to the distance *de*. That is, individual *A* is willing to pay *df* of q_2 to obtain Δq_1, whereas individual *B* requires only *ef* of q_2 to give up Δq_1. If this excess of willingness-to-pay over willingness-to-accept is not paid to individual *B*, and individual *B* is paid only the minimum amount necessary, then point *e* is Pareto superior to point *c* because individual *A* is made better off and individual *B* is no worse off. If the excess is paid to individual *B*, the movement is to point *d*, which is again Pareto superior to point *c* because individual *B* is made better off and individual *A* is no worse off. If any nontrivial portion of the excess is paid to individual *B*, both are made better off. Thus, all points, including end points, on the line *de* are Pareto superior to point *c*. Similar reasoning suggests that a movement from point *c* to any point in the lightly shaded area can be shown to be an improvement on the basis of the Pareto criterion. Thus, all points in the lightly shaded area are Pareto superior to point *c*.

Now consider comparison of point *c* with any point in the heavily shaded areas. All points in the heavily shaded areas in Figure 2.1 are on indifference curves that are lower for both individuals *A* and *B*. Hence, points in the heavily shaded regions are Pareto inferior to point *c* because at least one individual is worse off and neither individual is better off.

Finally, consider comparison of point *c* with all remaining points that are not in shaded areas in the Edgeworth–Bowley box to discover a major shortcoming of the Pareto criterion: at all these points, one person is made better off and the other person is worse off relative to point *c*. That is, at point *a*, individual *B* is better off than at point *c* but individual *A* is worse off. Hence, these points are noncomparable using the Pareto principle. *Improvements for society using the Pareto criterion can be identified only for cases where everyone gains or at least no one loses.*

Now suppose that society starts at point *c* and moves to point *e*, making a Pareto improvement. At point *e*, $\mathrm{MRS}^A_{q_1q_2} > \mathrm{MRS}^B_{q_1q_2}$ and, hence, further gains from trade are possible. Suppose that individuals *A* and *B* continue trading, with individual *A* giving up each time the minimum amount of q_2 necessary to obtain additional units of q_1. In this manner, the trade point moves along the indifference curve U_1^B until point *b* is reached. At point *b*, the amount of q_2 that individual *A* is willing to give up to obtain an additional unit of q_1 is just equal to the amount of q_2 that individual *B* would demand to give up a unit of q_1. With any further movement, individual *A* would not be willing to pay the price that individual *B* would demand. In fact, a movement in any direction from point *b* must make at least one person worse off. Thus, point *b* is a Pareto optimum.

In this manner, one can verify that the marginal condition that holds at point *b*,

$$\mathrm{MRS}^{A}_{q_1 q_2} = \mathrm{MRS}^{B}_{q_1 q_2}, \tag{2.1}$$

holds at all points on the contract curve. Thus, *in the pure exchange case, any point on the contract curve is a Pareto optimum. Pareto optimality implies that the marginal rate of substitution between any two goods is the same for all consumers.* The intuition of this condition is clear because improvements for both individuals are possible (and Pareto optimality does not hold) if one individual is willing to give up more of one good to get one unit of another good than another individual is willing to accept to give up the one unit.

2.3 PRODUCTION EFFICIENCY

Efficiency in production must also be considered when discussing Pareto optimality. Consider once again Figure 2.1 – but now assume that more of good q_1, good q_2, or both can be made available by improving the efficiency with which inputs are used. This would imply that individual *B*'s origin O_B could be moved rightward, upward, or both. In any of these cases, any two indifference curves that were previously tangent on the contract curve would now be separated by a region such as the lens-shaped, lightly shaded region in Figure 2.1. Thus, individual *A*, individual *B*, or both could be made better off. If production possibilities are considered and more of q_1, q_2, or both can be produced, then the points on the contract curve in Figure 2.1 will no longer be Pareto-optimal points. Stated conversely, where alternative production possibilities exist, the points on the contract curve $O_A O_B$ can be Pareto efficient only if the point $(\bar{q}_1, \bar{q}_2)$ is an efficient output bundle. A *Pareto-efficient output bundle* is one in which more of one good cannot be produced without producing less of another.

However, an output point can be efficient only if inputs are allocated to their most efficient uses. To see this, consider the production-efficiency frontier in the Edgeworth–Bowley box in Figure 2.2. This box is constructed by drawing the isoquant map for output q_1 as usual with isoquants q_1^1, q_1^2, and q_1^3, but with the isoquant map for q_2 upside down and reversed with origin at O_{q_2}. The total amounts of inputs x_1 and x_2 available are given by $\bar{x}_1$ and $\bar{x}_2$. Any point in this box represents an allocation of inputs to the two production processes. For example, at point *g*, x_1^1 of x_1 and x_2^1 of x_2 are allocated to the production of q_1. The remainder of inputs $\bar{x}_1 - x_1^1$ and $\bar{x}_2 - x_2^1$ are allocated to the production of q_2.

Point *g* does not represent an efficient allocation of inputs because at point *g* the *rate of technical substitution* of x_1 for x_2 in the production of q_1, denoted by $\mathrm{RTS}^{q_2}_{x_1 x_2}$, is greater than the *rate of technical substitution* of x_1 for x_2 in the production of q_2, denoted by $\mathrm{RTS}^{q_1}_{x_1 x_2}$. The rate of technical substitution measures the rate at which one input can be substituted for another while maintaining the same level of output. Thus, if an increment of x_1, say Δx_1, is shifted from the production of q_2 to q_1, then an increment of x_2, say Δx_2, could be shifted to the production of q_2 without decreasing the output of q_1 from q_1^1. But only an increment of $ab < \Delta x_2$ of x_2 is necessary to maintain the output of q_2 at the original level q_2^1. This results from the fact that the marginal rate at which x_1 substitutes for x_2 in q_2 is less than the rate at which it substitutes for x_2 in the production of q_1. If, in the exchange of Δx_1 for an increment of x_2, the output of q_2 is kept constant at q_2^1, the output

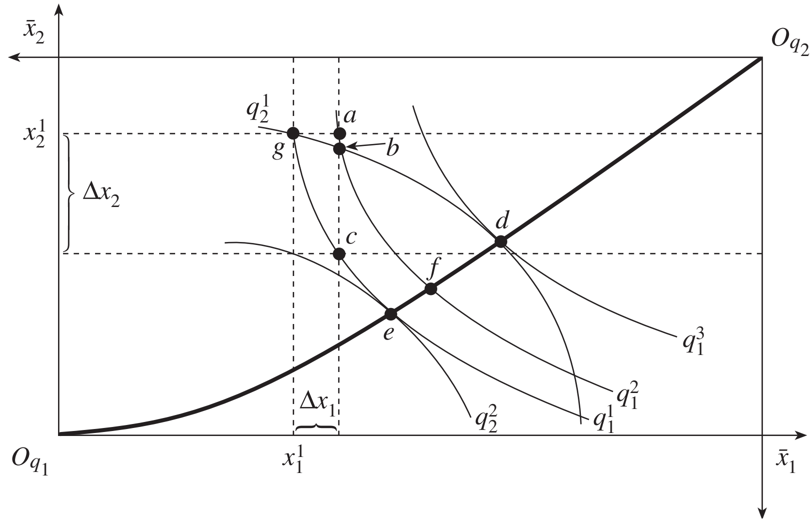

Figure 2.2

of q_1 can be increased to q_1^2. Thus, point *g* is an inefficient point. The output of q_2 can be increased without decreasing the output of q_1. Clearly, if any amount of x_2 on the segment *bc* is allocated to q_2 as Δx_1 is allocated to q_1, the outputs of both q_1 and q_2 are increased.

By identical reasoning, point *b* can be established as an inefficient point. The output of q_1 can again be increased, holding q_2 constant, by reallocating x_1 to q_1 and x_2 to q_2. This process can be continued until a state such as point *d* is reached. At point *d*, the amount of x_2 that can be given up in exchange for an increment of x_1, keeping q_1 constant at q_1^3, is precisely equal to the amount of x_2 needed to keep q_2 constant at q_2^1 if an increment of x_1 is removed from the production of q_2. It is impossible to increase the output of q_1 without decreasing the output of q_2. Point *d* is thus a Pareto-efficient output point. But it is not unique. For example, point *e*, a point of tangency between q_1^1 and q_2^2, is also a Pareto-efficient output point, as is point *f*. In fact, all points on the *efficiency locus* $O_{q_1}O_{q_2}$ are Pareto-efficient output points.

Tangency of the isoquants in Figure 2.2 implies that

$$\mathrm{RTS}^{q_1}_{x_1x_2} = \mathrm{RTS}^{q_2}_{x_1x_2}. \tag{2.2}$$

Thus, to the earlier exchange conditions, this second set of conditions for Pareto optimality can now be added. That is, *Pareto optimality in production implies that the rate of technical substitution between any two inputs is the same for all industries that use both inputs.* The intuition of this condition is clear because greater production of both goods is possible (and Pareto optimality does not hold) if one production process can give up more of one input in exchange for one unit of another input than another production process requires to give up that one unit (holding the quantities produced constant in each case).

The set of Pareto-optimal points (or the efficiency locus) $O_{q_1}O_{q_2}$ can also be represented in output space. In Figure 2.3, the curve connecting q_1^* and q_2^* corresponds to $O_{q_1}O_{q_2}$. That is, q_2^* is the maximum output possible if all factors, $\bar{x}_1$ and $\bar{x}_2$, are used in the production of q_2. This point corresponds to point O_{q_1} in Figure 2.2. Likewise, q_1^* in Figure 2.3 corresponds to point O_{q_2} in Figure 2.2. Similarly, one can trace out the entire set of production possibilities corresponding to the contract curve $O_{q_1}O_{q_2}$ in Figure 2.2. This efficiency locus in output space is called the *production possibility curve or frontier*. Thus *all Pareto-efficient production points are on the production possibility frontier*.

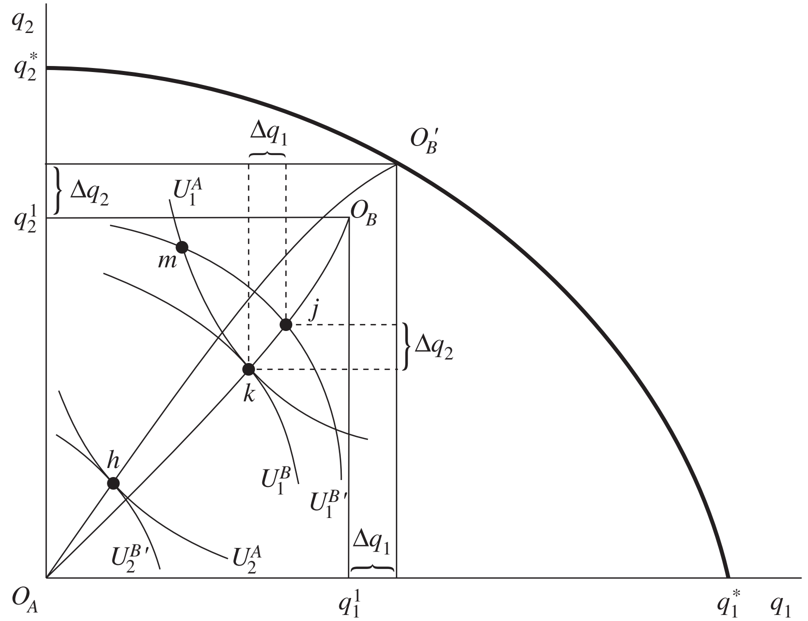

Figure 2.3

The Pareto inferiority of point *g* in Figure 2.2 is clear in Figure 2.3. The Edgeworth exchange box corresponding to output point *g* is drawn with individual *A*'s origin at O_A and individual *B*'s origin at O_B, with the output associated with point *g* in Figure 2.2 efficiently distributed at point *k*. A movement to point *f* in Figure 2.2 corresponds to providing the increments Δq_1 and Δq_2 in Figure 2.3, which shifts individual *B*'s origin to O_B'. Thus, individual *B*'s initial indifference curve U_1^B shifts to position $U_1^{B\prime}$, and his or her initial consumption point shifts to point *j*. The increase in output, $\Delta q_1 + \Delta q_2$, can then be used to make individual *A*, individual *B*, or both better off. Hence, point *g* in Figure 2.2 is not a Pareto optimum.

Although point *g* in Figure 2.2 is not a Pareto optimum, one cannot conclude that a movement from point *g* to any Pareto-efficient production point is a Pareto improvement. That is, the Pareto criterion cannot be used to compare all inefficient production points with all efficient ones. For example, a movement from O_B to O_B' in Figure 2.3 can be accompanied by a distribution of the larger output at point *h* which, although an efficient exchange and production point, results in individual *A* being made worse off and individual *B* being made better off than if output O_B is distributed at point *m*. Without a priori knowledge of the distribution of a larger bundle of goods, one cannot say whether or not a Pareto improvement occurs. *A Pareto improvement with a larger bundle of goods is attained only if it is distributed such that all individuals are made better off or no one is made worse off*. Thus, using the Pareto criterion, society cannot choose between states with more goods and states with fewer goods unless distributional information is also available.

2.4 THE PRODUCT-MIX CASE

From the foregoing, it should not be surprising that the Pareto criterion cannot be used to rank bundles of goods where one bundle has more of one good but less of another good without knowledge of how the goods are distributed. In what follows, this point is demonstrated, and the notion of Pareto optimality is discussed for the more general case where society has a choice over product mix.

First, however, the concept of the *Scitovsky indifference curve* (SIC) will be introduced. In Figure 2.4, society produces q_1^1 of good q_1 and q_2^1 of good q_2. The SIC labeled C in Figure 2.4 corresponds to point a on the contract curve, where the level of utility is represented by U_1^A for individual A and by U_1^B for individual B. To determine the SIC, hold O_A stationary, thus holding individual A's indifference curve stationary at U_1^A. At the same

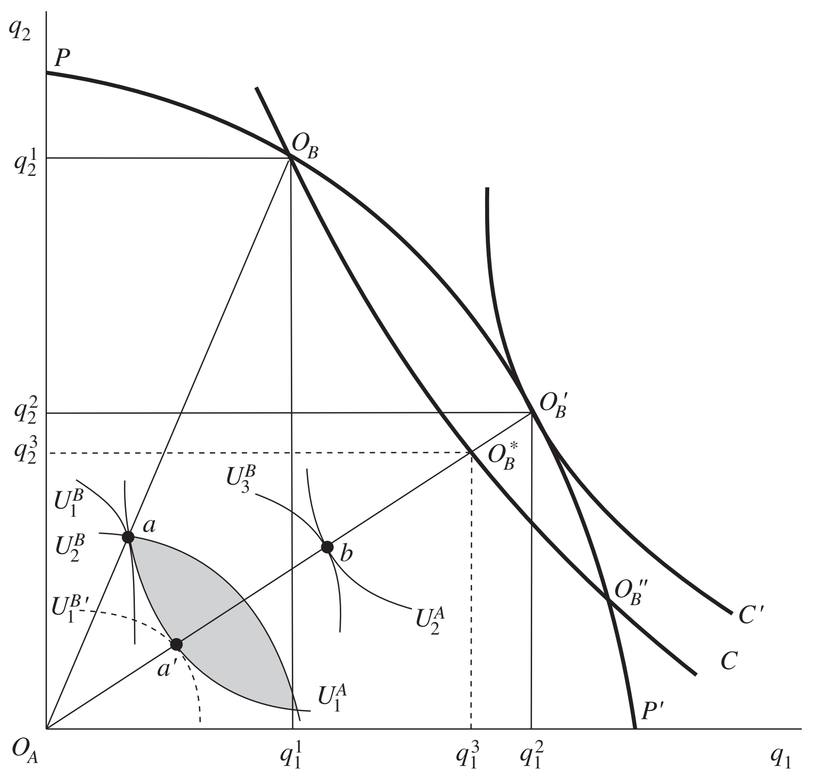

Figure 2.4

time, move O_B, thus shifting B's indifference map; but do this such that individual B's indifference curve U_1^B remains tangent to individual A's indifference curve U_1^A with axes for both goods parallel between individuals. Thus, C consists of the locus of all output bundles to which every member of society is indifferent if the bundle is initially distributed at point a.

Now consider comparing a particular distribution of one bundle of goods on the production possibility frontier with the distribution of another bundle also on the frontier. For example, in Figure 2.4 consider comparing the output bundle at O_B, which is distributed between individuals A and B at point a, with the output bundle at O_B' distributed at point a'. The Scitovsky curves corresponding to the distribution at points a and a' are C and C', respectively. Note that C is not tangent to the production possibility curve at O_B while C' is tangent at O_B'. Both individuals can be made better off by choosing the bundle at O_B' because C' lies above C and because only output at O_B^* (instead of O_B') distributed at point a' is needed to yield the same level of total utility as at O_B. The additional product $q_1^2 - q_1^3$ of q_1 and $q_2^2 - q_2^3$ of q_2 can be divided in any way desired to make both individuals better off in moving from O_B^* to O_B'.

Even though C' lies entirely above C, however, both individuals need not be made actually better off in moving from O_B to O_B' in Figure 2.4. That is, the bundle represented by O_B' may be distributed at point b, where individual A is made better off and individual B is worse off, relative to the bundle represented by O_B, distributed at point a. The SICs can lie with one entirely above the other, and one individual may still be worse off at the higher SIC. However, it is possible to redistribute the output bundle at O_B' to make everyone better off than at point a (the distribution of the initial bundle) by choosing a distribution of the product at O_B' in the shaded region.

Now consider comparing output bundle O_B' distributed at point a' (which generates the SIC denoted by C' tangent to PP') with any other efficient output bundle with all possible distributions. Because there are no feasible production points above C', it is impossible to generate an SIC that lies above C'. That is, if one starts at O_B' distributed at point a', one person cannot be made better off without making another person worse off. In other words, O_B' distributed at point a' is a Pareto-optimal point.

The requirement of tangency of the SIC to the production possibility curve thus establishes a third set of marginal equivalences which must hold for Pareto optimality in the product-mix case. That is, the slope of the production frontier must be the same as the slope of the SIC at the optimum. But the negative of the slope of the production possibility curve is the *marginal rate of transformation* of q_1 for q_2 (which measures the rate at which one output can be traded for another with given quantities of inputs), denoted by $\text{MRT}_{q_1q_2}$, and the negative of the slope of the SIC is the marginal rate of substitution of q_1 for q_2 for both individuals A and B. Thus, *Pareto optimality in product mix implies that the marginal rate of transformation must be equal to the marginal rates of substitution for consumers*; that is,

$$\text{MRT}_{q_1q_2} = \text{MRS}^A_{q_1q_2} = \text{MRS}^B_{q_1q_2}. \tag{2.3}$$

The intuition of this condition is clear because improvements for one individual are possible without affecting any other individual if production possibilities are such that the incremental amount of one output that can be produced in place of one (marginal) unit of another output is greater than the amount that some individual is willing to accept in place of that one unit. Of course, this condition does not define a Pareto optimum uniquely. *Any* point on the production possibility frontier distributed such that the corresponding SIC is tangent to the production possibility curve satisfies the condition. And, as in the pure exchange case, the Pareto criterion does not provide a basis for choosing among these points.

2.5 PARETO OPTIMALITY AND COMPETITIVE EQUILIBRIUM

Pareto optimality has thus far been examined independent of societies' institutional arrangements for organizing economic activity. However, fundamental relationships exist between the notion of Pareto optimality and the competitive market system as a mechanism for determining production, consumption, and the distribution of commodities. In particular, when a competitive equilibrium exists, it will achieve Pareto optimality. Moreover, if producers and consumers behave competitively, any Pareto optimum can be achieved by choosing an appropriate initial income distribution and appropriate price vector.

Before these relationships can be demonstrated, the concept of competitive equilibrium for a market system must be defined. Suppose that the economy consists of N traded goods, J utility-maximizing consumers and K profit-maximizing producers. Also, suppose that consumers and producers act competitively, taking prices as given. Let the demands by consumer j follow by $\boldsymbol{q^j} = \boldsymbol{\tilde{q}^j}(\boldsymbol{p}, m^j) = [\tilde{q}_1^j(\boldsymbol{p}, m^j),...,\tilde{q}_N^j(\boldsymbol{p}, m^j)]$, which represents a vector of quantities demanded of all goods by consumer j where $\boldsymbol{p} = (p_1,...,p_N)$ is a vector of prices for all goods and m^j is the income level of consumer j. In addition, let the supplies of consumer goods by producer k be represented by $\boldsymbol{q^k} = \boldsymbol{\hat{q}^k}(\boldsymbol{p}, \boldsymbol{w})$ and let the demands for factor inputs by producer k be represented by $\boldsymbol{x^k} = \boldsymbol{\hat{x}^k}(\boldsymbol{p}, \boldsymbol{w}) = [\hat{x}_1^k(\boldsymbol{p}, \boldsymbol{w}),...,\hat{x}_L^k(\boldsymbol{p}, \boldsymbol{w})]$ where $\boldsymbol{w} = (w_1,...,w_L)$ is a vector of all input prices. Finally, suppose factor ownership is distributed among consumers so that each consumer holds a vector of factor endowments $\boldsymbol{\tilde{x}^j} = (\tilde{x}_1{}^j,...,\tilde{x}_L{}^j)$ and thus has income $m^j = \Sigma_{l=1}^L w_l \tilde{x}_l^j$. Then suppose there exist vectors of prices $\boldsymbol{\bar{p}}$ and $\boldsymbol{\bar{w}}$ such that the sum of quantities demanded is equal to the sum of quantities supplied in all markets,

$$\sum_{j=1}^{J} \boldsymbol{\tilde{q}^j}(\boldsymbol{\bar{p}},m^j) = \sum_{k=1}^{K} \boldsymbol{\hat{q}^k}(\boldsymbol{\bar{p}},\boldsymbol{\bar{w}}),$$

$$\sum_{k=1}^{K} \boldsymbol{\hat{x}^k}(\boldsymbol{\bar{p}},\boldsymbol{\bar{w}}) = \sum_{j=1}^{J} \boldsymbol{\tilde{x}^j}.$$

The set of prices $\boldsymbol{\bar{p}}$ and $\boldsymbol{\bar{w}}$ then gives a competitive equilibrium.[1] Thus, *a competitive equilibrium is simply a set of prices such that all markets clear.*

1. This definition of a competitive equilibrium assumes free entry so that profits of firms are driven to zero. If profits are nonzero, then all profits must be distributed to consumers so that consumer j has income $m^j = \Sigma_{l=1}^L w_l \tilde{x}_l^j + \Sigma_{k=1}^K s_{jk}$ where s_{jk} is the share of producer k profit received by consumer j such that $\Sigma_{j=1}^J s_{jk} = \pi_k$ where π_k is the profit of firm k,

$$\pi_k = \Sigma_{n=1}^N p_n \hat{q}_n^k - \Sigma_{l=1}^L w_l \hat{x}_l^k.$$

A competitive equilibrium can be shown to exist if (1) all consumers have preferences that can be represented by indifference curves that are convex to the origin, and (2) if no increasing returns exist for any firm over a range of output that is large relative to the market.[2] Of course, *many competitive equilibria may exist depending upon the distribution of factor ownership or consumer income.*

The First Optimality Theorem

The first important relationship between competitive equilibrium and Pareto optimality is that, *when a competitive equilibrium exists, it attains Pareto optimality.*[3] This result, formally known as the first optimality theorem, is sometimes called the *invisible hand* theorem of Adam Smith (1937). In the *Wealth of Nations*, first published in 1776, Smith argued that consumers acting selfishly to maximize utility and producers concerned only with profits attain a best possible state of affairs for society, given its limited resources, without necessarily intending to do so. Although more than one best (Pareto-efficient) state of affairs generally exists, Smith was essentially correct.

To see this, first consider the case of consumer A displayed in Figure 2.5. To maximize utility, given the *budget constraint* II' associated with income m, consumer A chooses the consumption bundle $(\bar{q}_1,\bar{q}_2)$ which allows him or her to reach the highest possible indifference curve. Thus, the consumer chooses the point of tangency between II' and the indifference curve $\bar{U}^A$. At this tangency, $\mathrm{MRS}^A_{q_1q_2}=\bar{p}_1/\bar{p}_2$ because the former is the negative of the slope of the indifference curve and the latter is the negative of the slope of the budget constraint. But under perfectly competitive conditions, all consumers face the same prices. Thus, $\mathrm{MRS}^B_{q_1q_2}=\bar{p}_1/\bar{p}_2$ for consumer B and, hence,

$$\mathrm{MRS}^A_{q_1q_2}=\mathrm{MRS}^B_{q_1q_2}, \tag{2.4}$$

2. The first condition is a standard assumption of economic theory and needs no further comment. The problem that arises with increasing returns is that the average cost curve for the firm is continuously decreasing and the marginal cost curve is always below the average cost curve. With falling average costs, if it pays the firm to operate at all, then it pays the firm to expand its scale of operations indefinitely as long as output price is unaffected because the marginal revenue is greater than the marginal cost on each additional unit produced. If increasing returns exist over a large range of output, the percentage of the industry output produced by such a firm eventually reaches sufficient size to have an influence on price, and thus the firm will no longer be competitive. Hence, no profit-maximizing equilibrium exists for competitive firms in this case. As long as increasing returns are small, on the other hand, a competitive industry will consist of a great number of firms with the usual U-shaped average cost curves, and all of these profit-maximizing firms will operate at either the minimum or on the increasing portion of their average cost curve. For a rigorous development of the problem of existence and uniqueness of competitive equilibrium, see Quirk and Saposnik (1968) or Arrow and Hahn (1971).
3. Formally, this result requires that (1) *firms are technologically independent* and (2) *consumers' preferences are independent*. The first assumption implies that the output of each firm depends only on the input-use decisions it makes, and not on the production or input decisions of other firms other than quantities traded at competitive prices. The latter assumption implies that the utility function for each consumer contains as variables only items over which the consumer has a choice and not those quantities chosen by other consumers or producers other than quantities traded at competitive prices. Assumptions (1) and (2) jointly imply that no externalities exist. For a rigorous proof of this result, see Quirk and Saposnik (1968) or Arrow and Hahn (1971). A detailed discussion is also given by Arrow (1970, pp. 59–73). It is possible for a competitive equilibrium to achieve Pareto optimality in the presence of externalities if efficient markets exist for all external effects. These possibilities are discussed in Chapter 13.

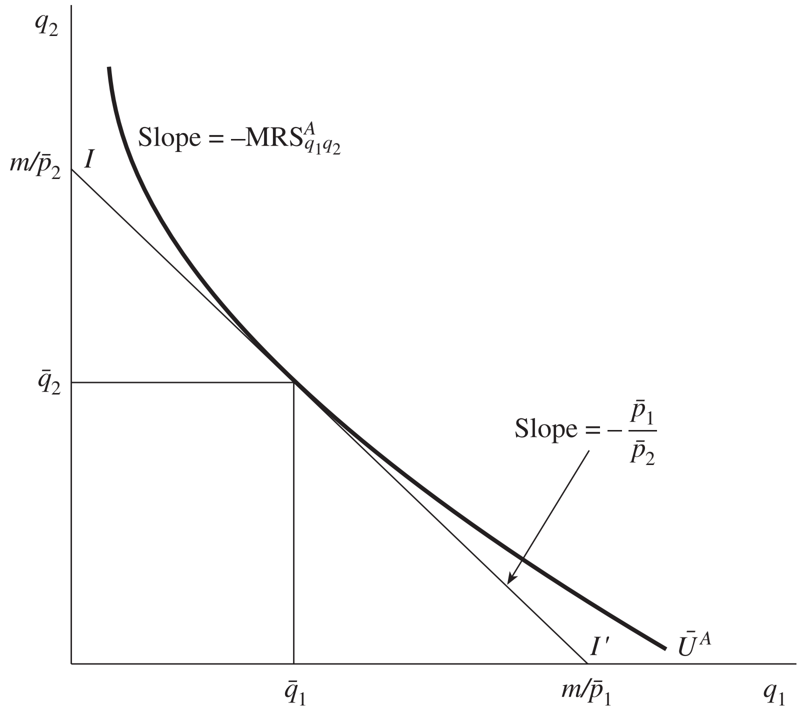

Figure 2.5

which is the Pareto-optimality exchange condition derived in equation (2.1). That is, because all consumers face the same relative prices of the two goods, their marginal evaluations must be the same in equilibrium.[4]

Now, recall that a firm cannot maximize profits for any level of output unless it is producing that output at a minimum cost. That is, profit maximization implies cost minimization. Assuming that $\bar{q}_1$ is the profit-maximizing level of output for the firm producing q_1 in Figure 2.6, the minimum cost of producing this output given input prices $\bar{w}_1$ and $\bar{w}_2$ is obtained by using $\bar{x}_1$ of x_1 and $\bar{x}_2$ of x_2. That is, the cost-minimizing input bundle is selected by finding the point of tangency of the *isocost curve CC′* (associated with cost level $\bar{c}$) with the isoquant associated with output $q_1 = \bar{q}_1$. Finding this point of tangency involves equating the slope of the isoquant, which is equal to the negative of the rate of technical substitution of x_1 for x_2, with the slope of the isocost curve, which is equal to

4. To develop this result more generally mathematically, let consumer j have utility function $U^j(\boldsymbol{q}^j)$ assumed to satisfy usual monotonicity, quasiconcavity, and differentiability properties. The consumer's budget constraint is then $m^j = \boldsymbol{pq}^j = \Sigma_{n=1}^N p_n q_n^j$, the Lagrangian of the utility maximization problem is $U^j(\boldsymbol{q}^j) - \lambda(m^j - \boldsymbol{pq}^j)$, and the first-order conditions are $\partial U^j/\partial q_n^j - \lambda p_n = 0$, $n = 1, \ldots, N$. Note that the demand functions $\boldsymbol{q}^j = \tilde{\boldsymbol{q}}^j(\boldsymbol{p}, m_j)$ must satisfy these first-order conditions. Taking ratios of pairs of first-order conditions implies that consumer behavior satisfies

$$\mathrm{MRS}^j_{q_n q_{n'}} \equiv \frac{\partial U^j/\partial q_n^j}{\partial U^j/\partial q_{n'}^j} = \frac{p_n}{p_{n'}}.$$

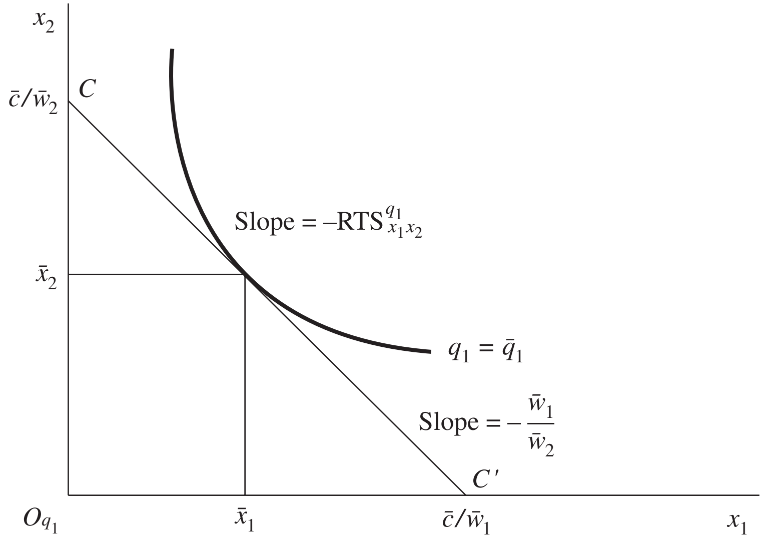

Figure 2.6

the negative of the ratio of input prices. But all producing firms that use x_1 and x_2 face the same input prices. Hence,

$$\mathrm{RTS}^{q_2}_{x_1 x_2} = \mathrm{RTS}^{q_1}_{x_1 x_2}, \tag{2.5}$$

which yields the production efficiency condition of equation (2.2).[5]

To establish that the product-mix condition of equation (2.3) is satisfied, first consider increasing the output of q_1 by Δq_1 and decreasing the output of q_2 by Δq_2 along the production possibility curve. Suppose this change is accomplished by transferring at the margin either one unit of x_1 or one unit of x_2 from production of q_2 to production of q_1. Then, the increase in output of q_1 is equal to the marginal physical product of input x_k in production of q_1. That is, $\Delta q_1 = \mathrm{MPP}^{q_1}_{x_k}$. Similarly, the decrease in output of q_2 is $\Delta q_2 = \mathrm{MPP}^{q_2}_{x_k}$. But the amount of q_2 that must be given up to obtain an increment of q_1 is given by the marginal rate of transformation between q_1 and q_2. Thus,[6]

5. To develop this result more generally mathematically, let producer k have short-run profit represented by $\pi^k = \boldsymbol{p}\boldsymbol{q}^k - \boldsymbol{w}\boldsymbol{x}^k = \Sigma_{n=1}^{N} p_n \hat{q}_n^k - \Sigma_{l=1}^{L} w_l x_l^k$, and implicit production function $f^k(\boldsymbol{q}^k, \boldsymbol{x}^k) = 0$ where f^k is assumed to satisfy usual monotonicity, concavity and differentiability properties. Then the Lagrangian of the profit maximization problem is $\boldsymbol{p}\boldsymbol{q}^k - \boldsymbol{w}\boldsymbol{x}^k + \mu[f^k(\boldsymbol{q}^k, \boldsymbol{x}^k)]$ and the first-order conditions are $\mu(\partial f^k/\partial q_n^k) + p_n = 0$, $n = 1, \ldots, N$, and $\mu(\partial f^k/\partial x_l^k) - w_l = 0$, $l = 1, \ldots, L$. Because all producer supplies and demands must satisfy these first-order conditions, taking appropriate ratios of pairs of first-order conditions implies

$$\mathrm{RTS}^k_{x_l x_{l'}} \equiv \frac{\partial f^k/\partial x_l^k}{\partial f^k/\partial x_{l'}^k} = \frac{w_l}{w_{l'}}.$$

6. The first-order conditions of footnote 5 also imply

$$\mathrm{MRT}^k_{q_n q_{n'}} \equiv \frac{\partial f^k/\partial q_{n'}^k}{\partial f^k/\partial q_n^k} = \frac{p_{n'}}{p_n},$$

which generalizes the result in equation (2.6).

$$\mathrm{MRT}_{q_1q_2} = \frac{\Delta q_2}{\Delta q_1} = \frac{\mathrm{MPP}_{x_k}^{q_2}}{\mathrm{MPP}_{x_k}^{q_1}}. \tag{2.6}$$

Now recall that cost minimization by a producer requires that producers equalize the marginal physical product per dollar spent on each input. That is, the least-cost combination $\bar{x}_1$, $\bar{x}_2$ in Figure 2.6 is characterized by the conditions

$$\frac{\mathrm{MPP}_{x_1}^{q_j}}{w_1} = \frac{\mathrm{MPP}_{x_2}^{q_j}}{w_2}, j = 1, 2. \tag{2.7}$$

But the marginal physical product of an input is equal to the increase in output Δq_j divided by the increase in input Δx_k required to obtain the increase in output. Thus, $\mathrm{MPP}_{x_k}^{q_j} = \Delta q_j / \Delta x_k$. Using this result in equation (2.7) and inverting yields

$$\bar{w}_1 \frac{\Delta x_1}{\Delta q_j} = \bar{w}_2 \frac{\Delta x_2}{\Delta q_j}, j = 1, 2, \tag{2.8}$$

where $\bar{w}_k\, \Delta x_k / \Delta q_j$ is simply the marginal cost, MC_{q_j}, of obtaining an additional unit of q_j. Combining (2.7) and (2.8) thus yields

$$\mathrm{MC}_{q_j} = \frac{\bar{w}_k}{\mathrm{MPP}_{x_k}^{q_j}} j = 1, 2. \tag{2.9}$$

Finally, recall that all profit-maximizing producers equate marginal revenue, which is simply the competitive producer's output price, with marginal cost. Thus, substituting $\bar{p}_j$ for MC_{q_j} in equation (2.9) and dividing the equation with $j = 1$ by the one with $j = 2$ yields[7]

$$\frac{\bar{p}_1}{\bar{p}_2} = \frac{\mathrm{MPP}_{x_k}^{q_2}}{\mathrm{MPP}_{x_k}^{q_1}}. \tag{2.10}$$

But from equation (2.6), the right-hand side of equation (2.10) is simply $\mathrm{MRT}_{q_1q_2}$. Because consumers face these same commodity prices, the product-mix condition in equation (2.3),

$$\mathrm{MRT}_{q_1q_2} = \mathrm{MRS}_{q_1q_2}^{A} = \mathrm{MRS}_{q_1q_2}^{B}, \tag{2.11}$$

must hold with competitive market equilibrium.

Thus, under the assumptions of this chapter, competition leads to conditions (2.4), (2.5), and (2.11), which are identical to conditions (2.1)–(2.3) that define a Pareto optimum. In other words, *competitive markets are Pareto efficient, meaning that competitive markets result in an equilibrium position from which it is impossible to make a change without making someone worse off.* This conclusion is probably the single most powerful

7. To obtain equation (2.10) from equation (2.6) generally, note that comparative static analysis of the production function constraint in footnote 5 holding all but one input and all but one output constant yields

$$\frac{dq_n^k}{dx_l^k} = -\frac{\partial f^k / \partial x_l^k}{\partial f^k / \partial q_n^k}.$$

Using this and a similar relationship for $dq_{n'}^k / dx_l^k$ in the equation of footnote 6 reveals that

$$\mathrm{MRT}_{q_nq_{n'}}^k = \frac{dq_n^k / dx_l^k}{dq_{n'}^k / dx_l^k} = \frac{\partial f^k / \partial q_{n'}^k}{\partial f^k / \partial q_n^k} = \frac{p_{n'}}{p_n}.$$

The result in equation (2.11) follows because each term is equal to the same price ratio.

result in the theory of market economies and is widely used by those economists who believe that markets are competitive and, hence, that government should not intervene in economic activity. Milton Friedman and the 'Chicago School' are the best known defenders of this position (see Friedman and Friedman 1980). In addition, because of its efficiency properties, competitive equilibrium offers a useful standard for policy analysis. For this purpose, states of competitive equilibrium or Pareto optimality are called *first-best states* and the associated allocations are called *first-best bundles*. All other states or bundles are called *second-best*. Departures from competitive equilibria are called *market failures*. Examples of market failures include monopolistic behavior, taxes, and externalities. Policies that correct market failures are thus viewed as achieving competitive equilibrium and therefore attain economic efficiency.

The Second Optimality Theorem

The second optimality theorem states that *any particular Pareto optimum can be achieved through competitive markets by simply prescribing an appropriate initial distribution of factor ownership and a price vector*.[8] That is, a central planner can achieve any efficient production bundle and any distribution of consumer well-being by redistributing factor ownership and prescribing appropriate prices where consumers maximize utility subject to budget constraints and producers maximize profits. The use of the competitive mechanism in this manner is sometimes called Lange–Lerner socialism after the two economists who first recognized this possibility (see Lange 1938; Lerner 1944). This result implies that many Pareto optima exist which are competitive equilibria, each associated with different factor endowments. The potential for widely differing marginal valuations under alternative competitive equilibria illustrates the connection between efficiency and income distribution.

Many economists object to addressing efficiency and distribution in two stages where the first stage involves maximizing economic efficiency and the second stage involves distributing the product equitably. The relative value of products depends on income distribution, which depends, in turn, on the factor ownership distribution. Actually, the Lange–Lerner result suggests the opposite approach whereby distributional objectives can be achieved by first redistributing factor ownership. Then policies need to be adopted only to correct market failures in order to achieve a Pareto optimum consistent with the desired income distribution.

Figure 2.7 demonstrates this point by considering only two possible states. The two goods produced are q_1 and q_2, and PP' is the production possibility frontier. The Scitovsky indifference curve C pertains to the output bundle O_B distributed among the individuals at point a. Alternatively, the output bundle O_B' distributed at point b yields the Scitovsky indifference curve C'. As points O_B and O_B' show, both bundles and their distributions lead to Pareto-optimal states. Thus, points O_B and O_B', with corresponding distributions at points a and b, respectively, are called *first-best states*, but neither is a unique optimum because the other is also an optimum in the same sense. For example, a factor ownership distribution that produces competitive equilibrium at point a may leave the economy poorly suited to achieve a distributional objective consistent with point b. On

8. For a rigorous proof of this result, see Quirk and Saposnik (1968) or Arrow and Hahn (1971).

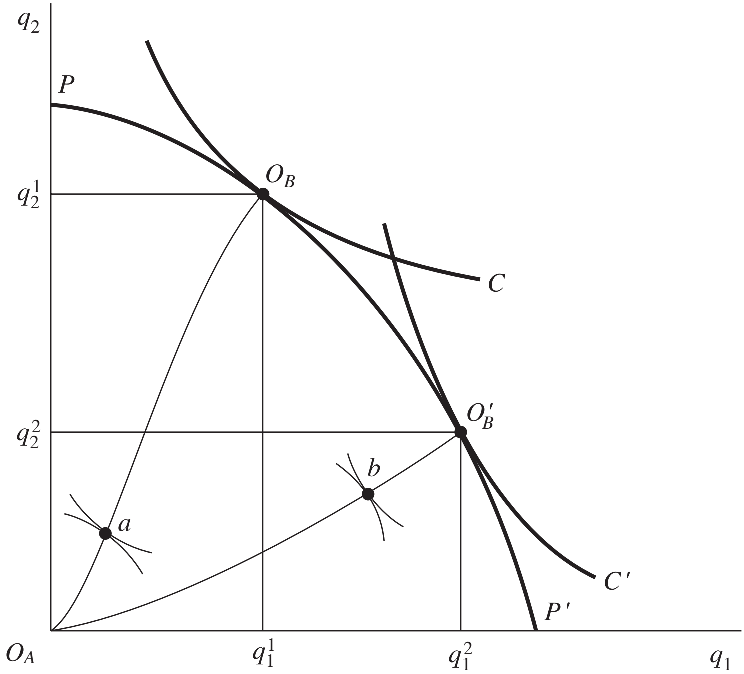

Figure 2.7

the other hand, starting from a factor ownership distribution that generates consumer incomes consistent with point *b*, the Lange–Lerner result implies that a Pareto efficient organization will be achieved automatically by the Adam Smith invisible hand in absence of market failures.

2.6 LIMITATIONS OF PARETO OPTIMALITY AND THE PARETO PRINCIPLE

Although the Pareto principle gives a plausible criterion for comparing different states of the world, its limitations are numerous. The greatest shortcoming of the Pareto principle is that many alternatives are simply not comparable. For example, in Figure 2.8, if production possibilities are represented by *PP* and production is initially O_B with distribution at point *b* corresponding to Scitovsky indifference curve *C*, then the only Pareto-preferred alternatives are in the shaded, lens-shaped area. All other production points are either infeasible, non-Pareto comparable or Pareto inferior. If production is initially at O_B with distribution at point *a* corresponding to Scitovsky indifference curve C', then no other feasible alternatives are Pareto superior. In fact, once *any* competitive equilibrium is reached in the framework of this chapter, no other feasible alternatives are Pareto superior. For example, production at O_B with distribution corresponding to the Scitovsky indifference curve C' is not Pareto comparable to production at O_B' with distribution corresponding to social indifference curve C^*. Thus, alternative Pareto optima are

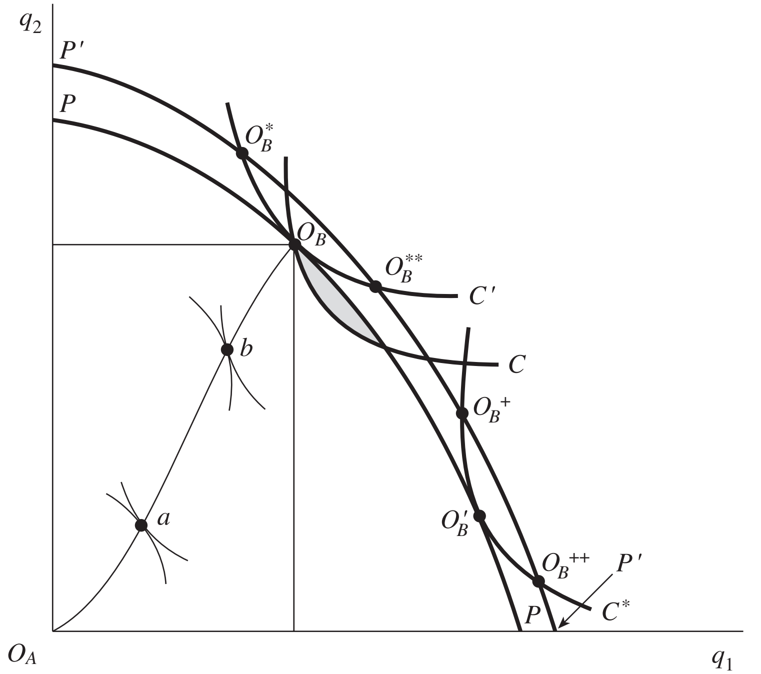

Figure 2.8

not Pareto comparable. Hence, the Pareto criterion prevents consideration of income-distributional considerations once a competitive equilibrium is attained.

Another serious problem with the Pareto principle is that no unique choice of distribution is apparent when improvements are possible. For example, suppose that a technological change takes place in Figure 2.8, shifting the production possibility frontier out to $P'P'$. If production was initially at O_B with distribution at point a corresponding to Scitovsky indifference curve C', the only points that are Pareto superior are those above C' and below or on $P'P'$. The only points that are Pareto superior and also possibly correspond to Pareto optimality with the new technology are those on the $P'P'$ frontier between $O_B{}^*$ and $O_B{}^{**}$. But the points along this interval may be associated with a wide variation in income distribution (assuming, for example, that only one distribution exists at each production point with tangency between the Scitovsky indifference curve and the new production possibility frontier). The Pareto criterion gives no basis for choosing among these alternatives. However, this problem may be viewed as an advantage because possibilities can exist for altering income distribution while fulfilling the simply appealing nature of the Pareto criterion. For example, in a rapidly growing economy (one with rapidly expanding production possibilities), the possibilities for altering income distribution while still fulfilling the Pareto criterion are many even though the Pareto criterion gives no guidance on which income distribution should be chosen.

Even with expanding production possibilities, however, the Pareto principle strongly favors the status quo. For example, if production is originally at O_B with distribution associated with Scitovsky indifference curve C', a new point of Pareto optimality is possible only at points between O_B^* and O_B^{**} on the new frontier $P'P'$ in Figure 2.8, assuming that the Pareto principle is satisfied in such a change. But if the initial production point is at $O_B{}'$ with distribution associated with Scitovsky indifference curve C^*, a new Pareto optimum is possible only at points between O_B^+ and O_B^{++} on the frontier $P'P'$, again assuming the Pareto principle is satisfied in the change. In each case the set of feasible Pareto improvements does not represent a substantial departure from the initial point unless technological improvements are large. Again, alternatives with widely varying income distributions may be neither comparable nor attainable from a given initial state by strict adherence to the Pareto criterion.

In a policy context, decisions often must be made where someone is made worse off while someone else is made better off. Furthermore, some policies are directly intended to change the distribution of income (that is, narrow the gap between high- and low-income people). Hence, to evaluate such changes, a device other than the Pareto principle is needed.

2.7 CONCLUSIONS

This chapter has focused on the concept of Pareto optimality and the Pareto criterion. A Pareto improvement is a situation where a move results in at least one person becoming better off without anyone becoming worse off. Pareto optimality is achieved when it is no longer possible for a policy change to make someone better off without making someone else worse off.

From a policy point of view, the Pareto criterion favors the status quo because the range of choices that represent Pareto improvements depends critically on the initial distribution of income. The Pareto criterion cannot be used to choose among widely different income distributions. Furthermore, many Pareto-optimal policy choices may exist that correspond simply to different income distributions. Perhaps not all first-best, Pareto-optimal choices are superior to some particular second-best choice. Although it is possible to make a Pareto improvement from a second-best state, it does not follow that any Pareto-optimal state is preferred to any second-best state. For example, if a second-best and a first-best state have markedly different income distributions, the situation that is second best may not be inferior to the first-best situation. Thus, the Pareto criterion alone appears to constitute an insufficient basis for applied economic welfare analysis of public policy alternatives.

3. The compensation principle and the welfare function

In Chapter 2, emphasis was given to the Pareto criterion as a means for selecting among alternative policies. Results show that many 'first-best' bundles or many Pareto-optimal points usually exist for an economy but, unfortunately, the Pareto principle does not give a basis for selecting among them. Narrowing the range of possibilities to a single first-best bundle (which essentially requires determining the ideal income distribution) requires a more complete criterion. One such criterion, which was introduced much later than the Pareto principle in the hope that it would be a more powerful device for choosing among policies, is the *compensation principle*, sometimes called the *Kaldor–Hicks compensation test* after the two economists to which it is attributed (Kaldor 1939; Hicks 1939). The development of the compensation principle was thus an attempt to broaden the states of the world that could be compared using an accepted welfare criterion. Simply stated, state B is preferred to state A if at least one individual could be made better off without making anyone worse off at state B – not that all individuals are *actually* no worse off – by some feasible redistribution following the change. Unlike the Pareto principle, the compensation criterion does not require the actual payment of compensation.

The issue of compensation payments is at the heart of many policy discussions. Some argue that compensation should be paid in certain cases. According to Lester Thurow (1980, p. 208), 'If we want a world with more rapid economic change, a good system of transitional aid to individuals that does not lock us into current actions or current institutions would be desirable.' However, most policies that have been introduced have not entailed compensation. For example, bans on DDT and other pesticides have in many cases resulted in producer losses, but producers have not been compensated for their losses in revenue.

Although the compensation principle does, in fact, expand the set of comparable alternatives (at the expense of additional controversy), some states remain noncomparable. The latter part of this chapter considers the necessary features of a criterion that ranks all possible states of an economy. However, empirical possibilities for the resulting more general theoretical constructs appear bleak.

3.1 THE COMPENSATION PRINCIPLE

According to the compensation principle, state B is preferred to state A if, in making the move from state A to state B, the gainers can compensate the losers such that at least one person is better off and no one is worse off. Such states are sometimes called *potentially Pareto preferred* states. The principle is stated in terms of *potential* compensation rather than *actual* compensation because, according to those who developed the principle, the payment of compensation involves a value judgment. That is, to say that society should move to state *B* and compensate losers is a clearly subjective matter, just as recommendation for change on the basis of the Pareto criterion is a subjective matter. For example, if a Pareto improvement is undertaken, then, as demonstrated in Section 2.6, the possibilities that represent further Pareto improvements may be more restricted. Conceivably, the true optimum state of society may not be reachable by further applications of the criterion if the wrong initial Pareto improvement is undertaken. Similarly, to say that society should move to state *B* without compensating losers is also a subjective matter of perhaps a more serious nature. Thus, nonpayment of compensation also involves a value judgment. In terms of objective practice, one can only point out the potential superiority of some state *B* without actually making a recommendation that the move be made.

The Pure Consumption Case[1]

Consider the application of the compensation principle to comparing different distributions of a given bundle of goods, again using the basic model of two goods and two individuals developed in Chapter 2. In Figure 3.1, point *a* is preferred to point *b* on the basis of the Pareto principle. But how does one compare point *b* with a point such as *c*, where *c* is not inside the lens-shaped area? The compensation principle offers one possibility. For example, suppose that one redistributes the bundle such that, instead of being at point *b*, individual *A* is at point *d* and individual *B* is at point *e*. Note that the welfare of each is unchanged. However, at these points there is an excess of q_2 equal to $q_2^3 - q_2^2$ and an excess of q_1 equal to $q_1^3 - q_1^2$. Now, if the move *actually* takes place to point *c*,

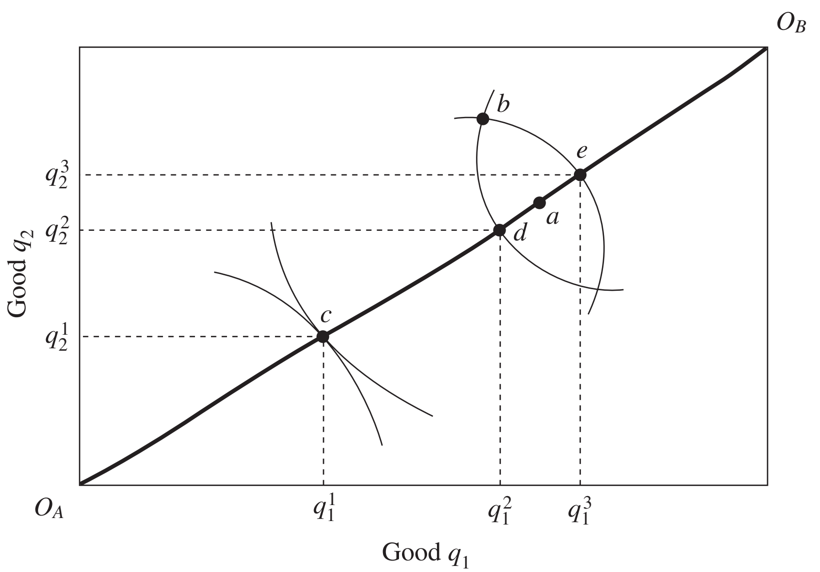

Figure 3.1

1. This section is largely based on Bailey (1954).

individual A is clearly worse off while individual B is better off. Individual A loses $q_2^2 - q_2^1$ of q_2 and $q_1^2 - q_1^1$ of q_1, but individual B gains $q_2^3 - q_2^1$ of q_2 and $q_1^3 - q_1^1$ of q_1. The amount individual B gains in physical amounts of q_1 and q_2 is greater than the loss to individual A. Hence, point c is potentially preferred to point b. *By the compensation principle, everyone is potentially better off by moving from point b to point c because the amount individual B gains is greater than the amount individual A loses.* This result holds even though compensation is not actually paid. If compensation is paid in terms of q_1 and q_2, both parties would, in effect, not agree to move to point c. Instead, a move would take place from point b only to somewhere within the lens-shaped area. But points within the lens-shaped area are comparable with point b by the Pareto principle. Thus, the application of the compensation criterion does not increase the ability to make statements about *actual* increases in welfare.

To view the problem in a different way, consider to what extent individual B would have to bribe individual A in order to make the move from point b to point c. The minimum amount is $q_2^2 - q_2^1$ of q_2 and $q_1^2 - q_1^1$ of q_1. Hence, in equilibrium, one would move from point b to point d only if compensation were paid. Individual B would gain $q_2^3 - q_2^2$ of q_2 and $q_1^3 - q_1^2$ of q_1 in the move if the minimum bribe is paid. Thus, point c is never actually reached if compensation is paid.

Distribution of Different Bundles

Consider next how the compensation principle can be used to compare different distributions of *different output bundles*. Recall from the preceding case that potential gains can be made in a move from point b to point c if, in the actual move to point c, the amount one individual loses is less than the amount the other individual gains. With this in mind, consider Figure 3.2 where the indifference curve C corresponds to production at O_B and to distribution at point a. Similarly, with production at O_B^*, the Scitovsky curve corresponding to distribution at point b is C^*. At point b, one individual is worse off than at point a, and the other individual is better off. However, potential gains are possible in the move from point a to point b because the amount the loser loses is less than what the gainer gains. Potential gains are clear because production at O_B^* can be distributed to keep welfare the same as at point a by moving along the Scitovsky indifference curve C to point f. By so doing, fh of q_2 and fg of q_1 are left over. Thus, if the compensation principle is used as a policy criterion, the move would be made (even though at point b one of the individuals may be actually worse off than at point a).

At this point, a comparison and contrast can be drawn between the compensation principle and the Pareto principle. Using the compensation principle with initial production bundle at O_B and distribution at point a, a move to the production bundle at O_B^* is supported regardless of the way it is actually distributed. Using the Pareto principle, however, the move is supported only if the actual distribution corresponds to moving along the Scitovsky curve C to point f, keeping the welfare of each individual constant and then dividing the excess of fg of good q_1 and fh of good q_2 among the two individuals in some way so that neither is worse off.

The reason that production at O_B^* is preferred to production at O_B, in either case, is that the starting point with distribution at point a is a second-best state. The corresponding Scitovsky curve C is not tangent to the production possibility frontier PP. Like the

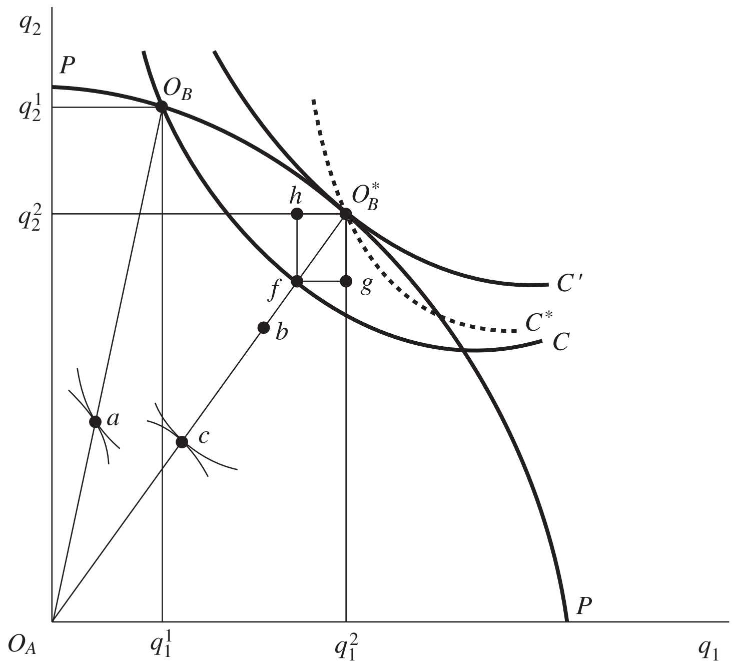

Figure 3.2

Pareto criterion, the compensation principle does not support a move away from a first-best state such as production at O_B^* with distribution point *c* corresponding to Scitovsky indifference curve C'. Thus, the compensation criterion, like the Pareto criterion, cannot be used to rank two first-best states. A movement from one to the other would not be supported regardless of which is used as a starting point. The compensation criterion, on the other hand, gives a means of comparing all pairs of second-best states and for comparing all second-best states with all first-best states.

The Reversal Paradox

An important class of problems in applying the compensation principle falls under the general heading of the *reversal paradox* pointed out by Scitovsky (1941). For the case where gainers can potentially compensate losers, a conclusion that one position is better than another is not always warranted. *One must ask, also, whether the losers can bribe the gainers not to make the move.* The crux of the argument is presented in Figure 3.3. The production possibility curve is *PP*, and the two bundles to be compared are O_B and O_B^*. Each of the bundles is distributed such that the corresponding Scitovsky indifference curves cross. In other words, both are second-best states because neither indifference curve is tangent to the production possibility curve. The curves C_1 and C_2 correspond to points *a* and *c* on the contract curves, respectively. Now, by the compensation principle, production at O_B^* with distribution at point *c* is better than production at O_B with distribution at point *a* because production at O_B^* can generally be redistributed such that all are actually

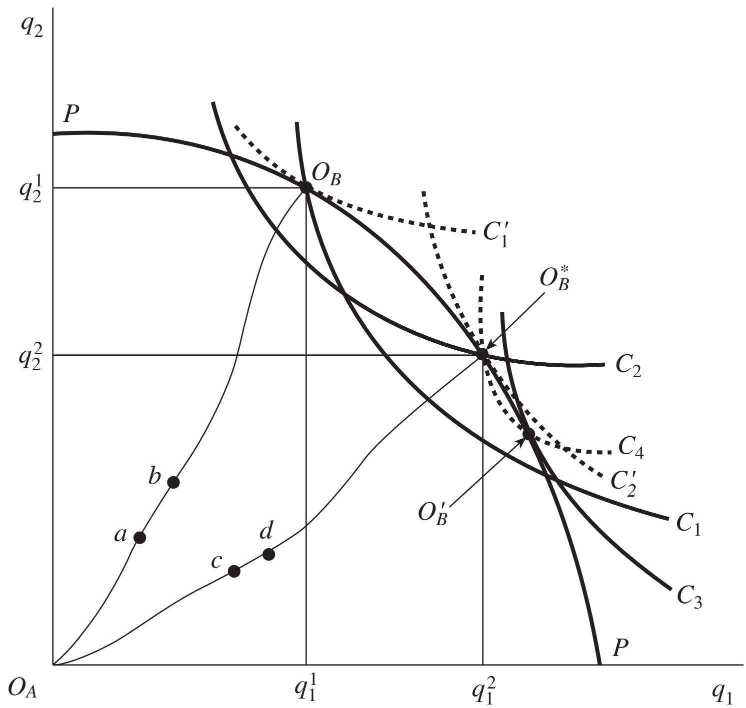

Figure 3.3

better off at point d (where distribution at point d corresponds to Scitovsky curve C_2', which lies above curve C_1 and is associated with improved welfare for both individuals). However, by this criterion, O_B^* is only potentially better off. Compensation is not actually paid. Because compensation is not actually paid, a reversal problem arises. That is, the new state with production at O_B^* and distribution at point c is a second-best state with Scitovsky curve C_2. Thus, according to Figure 3.3, there must be some distribution – say, at point b – such that production at O_B is preferred to production at O_B^* by the Scitovsky criterion (where distribution at point b corresponds to the Scitovsky curve C_1', which is associated with improved welfare for both individuals as compared with C_2). Thus, each is preferred to the other.

This reversal occurs because in each case a given distribution of the first bundle is compared with all possible distributions of the alternative bundle. The reversal paradox suggests that all distributions of the initial bundle should also be considered. In other words, a *reversal test* (sometimes called the *Scitovsky reversal test*) is passed if one determines, first, that gainers can bribe losers to make a change and, second, that losers cannot bribe gainers not to make the change. Unless the reversal test is passed in addition to the Kaldor–Hicks compensation test, one cannot really say that one state is even potentially preferred to another.

Some additional points that must be borne in mind with respect to the Scitovsky reversal paradox are as follows:

1. *The reversal paradox occurs only in comparing two second-best bundles*. It does not arise if one of the bundles is a first-best or Pareto-efficient bundle. For example, in Figure 3.3, if production at O_B^* with distribution corresponding to indifference curve C_2 is compared to production at O_B' with distribution corresponding to indifference curve C_3, a reversal problem does not occur.
2. *The reversal paradox does not always occur in comparing two second-best bundles even though compensation is not actually paid*. For example, in Figure 3.3, production at O_B^* with distribution corresponding to Scitovsky indifference curve C_4 does not lead to a paradox when compared to production at O_B and distribution corresponding to Scitovsky curve C_1. The paradox occurs only when the relevant Scitovsky curves cross in the interior of the feasible production region. This problem may not occur when income distributions do not change substantially.

Intransitive Rankings

If the compensation criteria (both the direct Kaldor–Hicks and Scitovsky reversal tests) are employed to rank all possible states, a further problem can arise even if the reversal problem is not encountered. That is, compensation tests can lead to intransitive welfare rankings when more than two states are compared.[2] This problem arises when, for example, one must choose among, say, *states where all the alternative policies are of a second-best nature* (that is, there is no single policy in the policy set that leads to a bundle of goods distributed with the Scitovsky community indifference curve tangent to the production possibility curve). In Figure 3.4, given the production possibility curve *PP*, bundle O_B^2 is preferred to O_B^1, O_B^3 is preferred to O_B^2 and bundle O_B^4 is preferred to O_B^3, using the compensation test. However, O_B^1 is also preferred to O_B^4. Hence, the Kaldor–Hicks compensation test leads to welfare rankings that are intransitive. But note that some form of distortion exists for each bundle because the Scitovsky indifference curves are not tangent to the production possibility curve in any of the four cases. All the bundles are of a second-best nature.

Suppose, on the other hand, that one policy results in a bundle of goods that is economically efficient (with the Scitovsky indifference curve tangent to the production possibility curve). For example, consider bundles O_B^1, O_B^2, O_B^3 and O_B^5. Here, O_B^5 is clearly the optimum choice. There is no desire, once at O_B^5, to return to bundles O_B^1, O_B^2 or O_B^3. As a second example, suppose that the bundles to be compared are O_B^1, O_B^2, O_B^3 and O_B^6. Again, once at O_B^6, no potential gain is generated in returning to O_B^1, O_B^2 or O_B^3. Hence, no ambiguity is encountered in choosing a top-ranked policy if the policy set contains exactly one first-best state. Thus, as with the reversal problem discussed earlier, intransitivity occurs only when all the bundles being compared are generated from second-best policies.[3]

Consider, on the other hand, one further case where the possibilities consist of O_B^1, O_B^5 and O_B^6. In this case, the Kaldor–Hicks compensation test shows that O_B^5 is preferred to

2. The results in this section are due to Gorman (1955).
3. Partly in response to the problems associated with using the compensation principle as a basis for welfare comparison, Arrow (1951) developed the impossibility theorem, which proves that no reasonable rule exists for combining rankings of various states of society by individuals into a societal ranking. See the further discussion in Section 3.4.

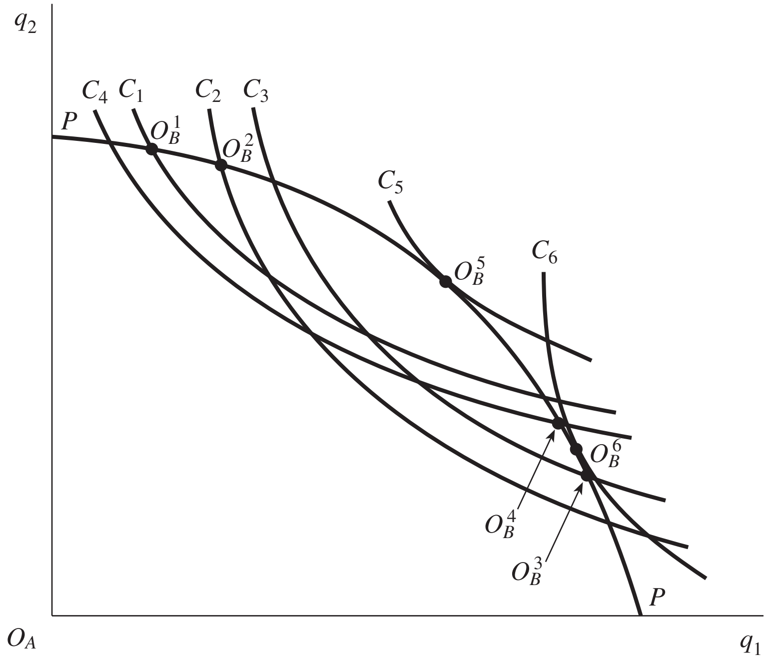

Figure 3.4

O_B^1. The possibility associated with O_B^6 is not preferred to O_B^1 even though O_B^6 is a first-best bundle. Among these three states, however, the rankings are not complete because, once at either O_B^5 or O_B^6, the compensation test does not suggest a move to either of the other states. In other words, the compensation test does not lead to a ranking of policy sets containing more than one first-best state.

3.2 UTILITY POSSIBILITY CURVES AND THE POTENTIAL WELFARE CRITERION

Another approach related to the choice of alternative income distributions and the reversal problem is based on the concept of utility possibility curves introduced by Samuelson (1947, 1956). To develop this approach, consider Figure 3.5 where the utilities of two individuals, *A* and *B*, are represented. The utility of individual *B* is measured on the vertical axis, while that for individual *A* is measured along the horizontal axis. Three utility possibility curves are represented, each of which is derived by changing the distribution of a given bundle of goods along a contract curve. For example, Q_2Q_2 shows the maximum utility both individuals can receive from a fixed production at O_B in Figure 3.3, Q_1Q_1 corresponds to a different bundle of goods, and so on.

To demonstrate the reversal paradox, consider Q_2Q_2 and Q_1Q_1. Points *a* and *b* represent particular distributions of the bundle from which Q_2Q_2 is derived. Similarly, points *c* and *d* represent particular distributions of the bundle from which Q_1Q_1 is derived.

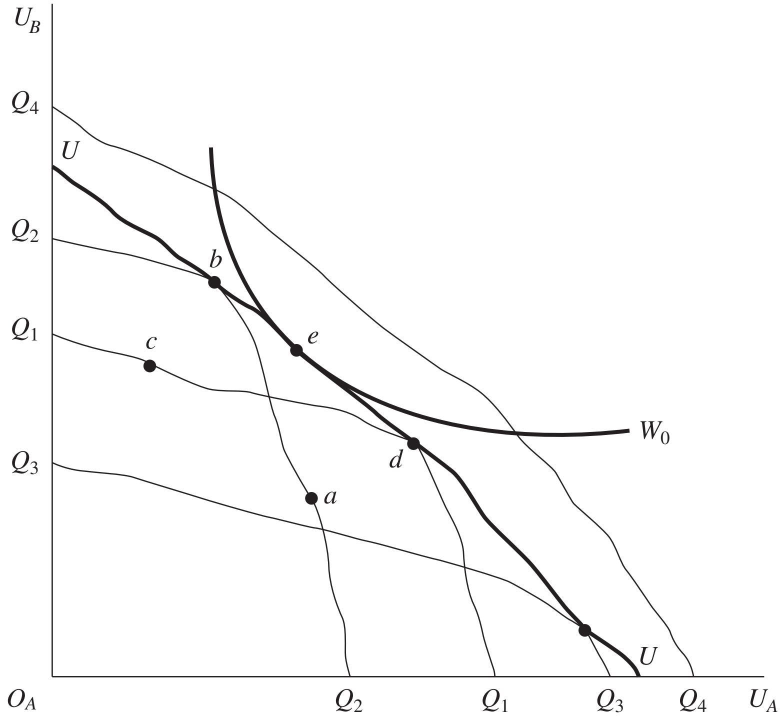

Figure 3.5

Suppose that the initial distribution is at point *c*. Then one can redistribute production when moving from Q_1Q_1 to Q_2Q_2 such that both individuals *A* and *B* would be better off at point *b* than at point *c*. Similarly, one could redistribute the other bundle so that both are better off at point *d* than at point *a*. The paradox arises because point *d* lies to the northeast of point *a*, while point *b* lies to the northeast of point *c*. Thus, one comparison implies a preference for the production bundle associated with Q_1Q_1, whereas the other comparison favors the production bundle associated with Q_2Q_2. This paradoxical situation would not arise if compensation were actually paid.

These results correspond directly with the analysis in commodity space in Figure 3.3. Points *a* and *c* in Figure 3.5 correspond to distributions that are second-best states. In other words, these points correspond to points *a* and *c* in Figure 3.3, which are also distributions giving rise to second-best states. Note that points *b* and *d* in Figure 3.5 and points *b* and *d* in Figure 3.3 correspond to first-best states.

If one considers all possible production bundles that can be obtained from a given production possibility frontier and all possible distributions of these bundles (which, in utility space, corresponds to considering utility possibility frontiers, such as Q_3Q_3, associated with all other possible production bundles), then the *grand utility possibility frontier UU* can be constructed as an envelope of the utility possibility frontiers. All points on this envelope curve correspond to first-best optima, that is, bundles distributed such that the Scitovsky curves are tangent to the production possibility curve.

Samuelson (1950) has argued that even if gainers can profitably bribe losers into accepting a movement and the losers cannot profitably bribe the gainers into rejecting it (that is, both the Kaldor–Hicks and Scitovsky criteria are satisfied), a potential gain in welfare is not necessarily attained. He argues that one has to consider all possible bundles and all possible distributions of these bundles before statements can be made about potential gains. The problem then amounts to selecting one among many first-best states for which there is no solution unless a social ranking of first-best utility possibilities can be determined. He proposes an alternative *potential welfare criterion*, which is demonstrated in Figure 3.5. Simply stated, *if there is some utility frontier such as Q_4Q_4 that lies entirely on or outside another utility frontier – Q_2Q_2, for example – owing perhaps to technological change, then any position on this new frontier is clearly at least potentially superior to any position on the old one.* Only if the new frontier lies entirely outside the other, however, are potential increases in real income necessarily obtained. Of course, this criterion can be used to compare either grand utility possibility frontiers before and after, say, a technological change or utility possibilities associated with given alternative production bundles.

In the absence of a rule for ranking alternative first-best utility possibilities, Samuelson (1950) argues that this is the only appropriate criterion to apply. In a strict sense, this argument is correct. But in a practical sense, this approach leads to few cases in which beneficial empirical evidence can be developed for policy-makers (see Section 8.3 for a discussion of the related empirical approach). On the other hand, the arguments in favor of this approach are based on an attempt to determine optimal policy without relying on policy-maker preferences or judgment. Such information is not necessary in a practical policy-making setting where political institutions exist for the express purpose of providing policy-makers to make such choices.

3.3 THE SOCIAL WELFARE FUNCTION

Because the potential welfare criterion often may not be satisfied even if utility possibility curves can be identified, economic inquiry has continued to search for a rule that can rank all states of society and thus determine which first-best state on the grand utility possibility frontier represents *the* social optimum. In theory, the social welfare function is such a concept. The *social welfare function* is simply a function – say, $W(U_A, U_B)$ – of the utility levels of all individuals such that a higher value of the function is preferred to a lower one. The assumption that the social welfare function is determined by the utilities of all individuals has been called the *fundamental ethical postulate* by Samuelson (1947) and is a cornerstone of democratic societies. Such a welfare function is called a *Bergsonian welfare function* after Abram Bergson (1938), who first used it.

The properties one would expect in such a social welfare function with respect to the utilities of individuals are much like those one would expect in an individual's utility function with respect to the quantities of commodities consumed. That is, one would expect that (1) an increase in the utility of any individual holding others constant increases social welfare (the Pareto principle); (2) if one individual is made worse off, then another individual must be made better off to retain the same level of social welfare; and (3) if some individual has a very high level of utility and another individual has a very low level of utility, then society is willing to give up some of the former individual's utility to obtain even a somewhat smaller increase in the latter individual's utility, with the intensity of this trade-off depending upon the degree of inequality.

The properties described above suggest the existence of welfare contours such as W_0 in Figure 3.5, which correspond conceptually to indifference curves for individual utility functions. By property (1), social welfare increases by moving to higher social welfare contours, either upward or to the right. By property (2), the social welfare contours have negative slope. By property (3), the welfare contours are convex to the origin.

Social welfare is maximized by moving to the highest attainable social welfare contour, which thus leads to tangency of the grand utility possibility frontier with the resulting social welfare contour such as at point *e* in Figure 3.5.[4] This tangency condition is sometimes called the *fourth optimality condition*. This condition, together with conditions in equations (2.1), (2.2) and (2.3), completely characterizes the social optimum.

3.4 LIMITATIONS OF THE SOCIAL WELFARE FUNCTION APPROACH

Although a social welfare function is a convenient and powerful concept in theory, its practical usefulness has been illusory. Many attempts have been made to specify a social welfare function sufficiently to facilitate empirical usefulness but none have been widely accepted. Apparently, little hope exists for determining a social welfare function on which general agreement can be reached. The major approaches that have been attempted include (1) the subjective approach, (2) the basic axiomatic approach and (3) the moral justice approach.

The subjective approach is represented by those who postulate a complete functional form for the social welfare function on subjective ethical grounds. Early students of the utilitarian school (for example, Bentham 1961, first published 1823) believed that changes in happiness should simply be added over individuals,

$$W = U^1 + U^2 + U^3 + \dots \tag{3.1}$$

4. Note that the slope of the welfare contour can be represented by

$$-\frac{\partial W / \partial U_A}{\partial W / \partial U_B} \equiv -\frac{W_{U_A}}{W_{U_B}}$$

if $W(U_A, U_B)$ is continuous and first derivatives exist. The slope of the utility possibility frontier is

$$-\frac{\partial U_B / \partial q_1}{\partial U_A / \partial q_1} = -\frac{\partial U_B / \partial q_2}{\partial U_A / \partial q_2}.$$

Thus, the tangency condition can be represented mathematically by

$$\frac{W_{U_A}}{W_{U_B}} = \frac{\partial U_B / \partial q_i}{\partial U_A / \partial q_i},\ i = 1, 2.$$

Cross-multiplying yields

$$\frac{\partial W}{\partial U_B}\left(\frac{\partial U_B}{\partial q_i}\right) = \frac{\partial W}{\partial U_A}\left(\frac{\partial U_A}{\partial q_i}\right),\ i = 1, 2,$$

which, simply stated, implies that the marginal social significance of consumption must be equated across individuals for each commodity.

where U^i represents the utility of individual i. A positive net gain is then viewed as grounds for policy implementation. This implies that the welfare contours such as W_0 in Figure 3.5 should be straight lines with slope −1. Others argue that a functional form should be used that reflects positive benefits from increases in equality consistent with 'normal' distributional judgments (see, for example, Blackorby and Donaldson 1990). One social welfare function that reflects inequality aversion is the form,

$$W = \frac{1}{1-\rho}[(U^1)^{1-\rho} + (U^2)^{1-\rho} + (U^3)^{1-\rho} + \ldots]. \tag{3.2}$$

The problem here is that the term 'normal' is ambiguous. For example, agreement cannot be reached on the appropriate level of inequality aversion, for example, the appropriate value of ρ in equation (3.2).[5]

The axiomatic approach, on the other hand, attempts to investigate the existence and form of the social welfare function mathematically based on a set of plausible underlying axioms about individual preferences and how they count to society. The most celebrated of these efforts is Arrow's (1951) *impossibility theorem*. This theorem addresses the question of whether a general rule exists that can rank social states based only on the way these states are ranked by individual members of society. Arrow showed that no such rule exists under the following plausible requirements:

1. *The domain of decisions is unrestricted.*
2. *The Pareto principle applies.*
3. *Dictatorship is ruled out.*
4. *Rankings are independent of irrelevant alternatives.*

An example of a rule that does not work is majority voting. Arrow's (1951) results suggest that social preferences are determined by a dictator (or a group that acts as a dictator), that the intensity of preferences of individuals rather than simple rankings matters (see Kemp and Ng 1977), or that one of the other axioms such as independence of irrelevant alternatives does not apply (see Sen 1970). Accordingly, Arrow's work has spawned a voluminous literature on possibility theorems by relaxing his axioms in various ways (see Sen 1982 or Fishburn 1973 for surveys). A major practical problem with this approach is that even under weaker axioms where voting works, the transactions costs of compiling votes or rankings of all individuals on each policy issue are prohibitive.

The moral justice approach argues that basic axiomatic examinations following Arrow fail because majority groups acting selfishly will prefer to eliminate consideration of minority interests. This failure can be addressed by admitting moral considerations such as *impartiality* and *economic justice*. Suppose that society consists of three individuals and a change is considered that takes $1000 from one individual to give $300 to each of the other two. If the three individuals were to vote selfishly knowing who the benefactors are, the majority would favor the change. On the other hand, if the voting were done not knowing who would pay and who would receive (a *veil of ignorance*), then the change

5. To demonstrate the different inequality aversions possible with the function in equation (3.2), note that it reduces to (3.1) when $\rho = 0$, it approaches (3.4) below when ρ approaches infinity, and it approaches the multiplicative form, $W = U^1U^2U^3\ldots$ when ρ approaches 1. See Boadway and Bruce (1984, ch. 5).

would be unanimously rejected. Alternatively, moral concerns for equal treatment of individuals (*impartiality*) have led some to support value judgments whereby the social welfare function treats individuals symmetrically, for example, $W(U^A, U^B) = W(U^B, U^A)$. The contours of the social welfare function in Figure 3.5 are then symmetric about a 45° line from the origin. If all individuals have identical utility functions then the utility possibility frontier is also symmetric about the 45° line and optimality is achieved by perfect equality. On the other hand, if one individual receives proportionally more utility from consuming the same bundles of goods as another, then such a welfare function would, in effect, assign different weights to the consumption of the individuals.

Harsanyi (1953, 1955) gave the first formal treatment of moral considerations by distinguishing between an individual's personal preferences and moral preferences. In his work, moral preferences are the rankings of a rational individual given that the individual does not know which set of personal preferences he or she will have. Under a relatively weak set of assumptions, Harsanyi (1953, 1955) shows with this approach that the social welfare function is a weighted sum of individual utilities,[6]

$$W = \alpha_1 U^1 + \alpha_2 U^2 + \alpha_3 U^3 + \ldots . \tag{3.3}$$

Further imposing impartiality (symmetry), the welfare function in (3.3) reduces to the Benthamite welfare function in (3.1) with equal weights. This welfare function has been called the *just social welfare function* (see Mueller 1979).

Other moral considerations, however, tend to suggest a stronger concept of equality. Moral considerations in economic welfare issues are often called *rights* to economic justice. Various value judgments or ethical postulates representing these moral considerations include the right to consume what one produces, the right to subsistence, the right to economic liberty and the right to economic equality.[7] With these considerations, taking \$1000 from a very rich individual to give \$300 to each of two poor individuals may be preferred on moral grounds. The most celebrated work in this area is Rawls's (1971) *Theory of Justice*. This theory, which is really more of a philosophy than a theory, contends that policy should be evaluated by the welfare of the most miserable person in society. This implies a social welfare function of the form

$$W = \min\,(U^1, U^2, U^3, \ldots). \tag{3.4}$$

In a more general framework, Arrow (1973) and Harsanyi (1975) show that this choice would be supported by individuals' moral preferences only under infinite risk aversion about the vested interests and preferences to be assumed. With other levels of risk aversion, the welfare function in (3.2) is obtained. Arrow concludes that the possibilities of discovering a theory of justice are remote given the diversity of ethical beliefs in society.

Virtually all of these moral consideration approaches suggest a criterion of distribu-

6. The Harsanyi assumptions are that both personal and moral preferences satisfy the von Neumann–Morgenstern axioms of choice, that each individual has an equal probability of taking on any individual's personal preferences, and that two states are socially indifferent if they are indifferent for every individual. Thus, choices are made according to expected utility given uncertainty about individual preferences.
7. For a more detailed review of the theories of ethical income distribution and economic justice, see Boadway and Bruce (1984, ch. 6).

tional optimality that tends, in some sense, toward either equality or equal weighting. However, even these two simple alternatives represented by (3.1) and (3.4) differ drastically in their implications. With (3.4), the worst-off individual becomes a dictator while, with (3.1), individuals who have very small utilities (and marginal utilities) tend not to matter. Nevertheless, each of these functions can be supported by a plausible set of axioms.[8] Thus, while axiomatic developments have added to the sophistication of social welfare function specification efforts, the effect has been to shift the level of disagreement from the function itself to the axioms that support it. Agreement on the set of axioms appears to be no more possible than agreement on the form of the social welfare function.

In summary, efforts to reach a unique social welfare function have not gained widespread acceptance in spite of great effort by a host of social choice theorists and moral philosophers. Thus, no generally acceptable or objective way to make interpersonal comparisons of utility exists.[9] In spite of the lack of agreement on form, adoption of specific alternative social welfare functions is still advocated from time to time in the literature. Some observed policy choices that strictly redistribute income by, for example, taxing the rich to give to the poor, cannot be advocated or explained with other economic criteria used for policy evaluation. Even if a social welfare function is determined, however, a host of practical problems arise in any practical application. The social welfare function approach requires that individual utilities are cardinally measurable so that intensities of preferences can be compared. In contrast to this approach, Pareto and compensation criteria assume only that utility can be measured ordinally. Thus, much greater practical applicability is attained even though the associated social ordering is not sufficiently complete to identify a unique social optimum or resolve questions of income distribution.

In applied welfare economics, the notion of a social welfare function is useful conceptually but one should keep in mind that a welfare function cannot be specified for practical purposes. However, this does not mean that the study of welfare economics is impractical because the function cannot be specified. Even those who are critical of welfare economics for this reason must agree that economists can make a useful contribution by pointing out who loses and who gains, as well as the magnitude of losses and gains caused for various groups by particular policies.

To summarize the welfare function controversy, it suffices to quote a notable welfare economist, E.J. Mishan (1973, pp. 747–8):

> The social welfare function, even when it is more narrowly defined as a ranking of all conceivable combinations of individual welfare, remains but a pleasing and nebulous abstraction. It cannot be translated into practical guidance for economic policy. Even if there were no fundamental obstacles to its construction, or even if one could think up reasonable conditions under which a social welfare function could exist, there would remain the virtually impossible task of arranging for society to rank unambiguously all conceivable combinations of the individual welfares and moreover – in order to utilise this massive apparatus – to discover (without much cost)

8. For example, Maskin (1978) and Sen (1982) find that any social welfare function with unrestricted domain that satisfies independence of irrelevant alternatives, the Pareto principle, anonymity, separability of unconcerned individuals, and cardinality with interpersonal comparability must be of the form in (3.1). However, simply replacing cardinality with ordinality and adding a minimal equity assumption (the best-off individual's preferences can never be served when they conflict with all worse-off individuals' preferences) results in the Rawlsian social welfare function in (3.4).
9. For further discussion of the difficulties related to determination of a social welfare function, see Atkinson (1970) and Sen (1973).

the effect on the welfare of each person in society (in terms of utilities, goods, or money) of the introduction of alternative economic organisations. For only if we have such data can we rank all existing and future economic possibilities and pronounce some to be socially superior to others. Although one can always claim that 'useful insights' have emerged from the attempts to construct theoretical social welfare functions, the belief that they can ever by translated into useful economic advice is chimerical.

In contrast, the more pedestrian welfare criteria, although analyzed in abstract terms, can be translated into practical propositions. Modern societies do seek to rank projects or policies by some criterion of economic efficiency and to take account also of distributional consequences.

3.5 POTENTIAL VERSUS ACTUAL GAINS

Because the social welfare function is a concept upon which general agreement has not been reached and because the potential welfare criterion is one that renders many policy alternatives noncomparable, the compensation principle has emerged as the criterion that is empirically the most widely applicable. But this state of affairs underscores the controversy about whether compensation should actually be paid when adopting policy changes that satisfy the criterion. If possible, should the gainers from a new policy actually compensate the losers so that 'everyone' is actually made better off? Should a policy change be recommended only on the basis of 'potential' gains alone, given that, if the change is made, someone is actually made worse off? As an example, the United Automobile Workers (UAW) union has taken the stand that new technology that displaces workers should not be introduced unless the workers are fully compensated for their losses. This is a case where the *potential* gains criterion is not supported. But to the extent that the UAW represents displaced workers, objections from the losing groups are not surprising.

However, an economist can often analyze the distributional impacts of policy choices without getting into the issue of compensation. For example, suppose one did an analysis of the impact of removing quotas on the importation of steel into the USA. A proper analysis would show the separate effects on government revenues, producers, consumers and the like (possibly by disaggregated groups if, say, several groups of consumers are affected differently). Thus, the losers, the gainers, and the magnitudes of losses and gains would be identified. Such an analysis would be useful to government officials who are elected or appointed to decide, among other things, the issue of compensation. In fact, a welfare analysis that does not adequately indicate individual group effects may be misleading or useless to government officials who have the authority to make interpersonal comparisons. Thus, as emphasized in Chapter 1, studies on the impact of policy choices can be done using welfare economics without getting into the debate as to what 'ought to be'.

3.6 PRACTICAL APPLIED POLICY ANALYSIS: THE RELATIONSHIP OF GENERAL EQUILIBRIUM AND PARTIAL EQUILIBRIUM ANALYSIS

The practical applicability of the various criteria for policy evaluation depends on the potential for empirical implementation and on the intuitive understandability of policymakers. Both the social welfare function and potential welfare criteria suffer in both respects. First, consumer utility cannot be measured sufficiently for empirical implementation under general conditions and, second, units of measurement for utility and social welfare are abstract and not well understood by policy-makers. The Pareto and compensation criteria, however, can be implemented in terms of individuals' *willingness to pay* and *willingness to accept* the effects of policies and projects. As demonstrated throughout this book, these measures can be reported in monetary terms that are generally empirically feasible and well understood.

A second problem with the framework used thus far for practical policy analysis is that an abstract general equilibrium framework has been used to investigate possibilities for identifying potential social gains through application of compensation criteria. Such a general equilibrium framework is highly useful for understanding the nature of problems encountered in application of compensation criteria, but it is not very helpful for analyzing and quantifying the implications of specific policies or projects involving markets and prices for specific goods. Policy-makers are generally concerned with impacts on specific markets and specific types of agents in society. The remainder of this book focuses on measuring individual, market and group-specific welfare effects by first concentrating on partial equilibrium models.

To facilitate the transition from general equilibrium analysis to the analysis of specific markets and agents, consider Figure 3.6. Figure 3.6(a) illustrates a production possibility curve PP, a Scitovsky indifference curve C, and a first-best equilibrium at (q_1^*, q_2^*), which attains a tangency of the production possibility curve and Scitovsky curve at prices p_1^* and p_2^* for goods q_1 and q_2, respectively. Figure 3.6(b) illustrates the supply and demand curves for q_1, which are derived from Figure 3.6(a) by varying the price p_1. The supply curve is found by plotting the absolute value of the slope of PP for each level of q_1. In other words, it is found by varying the price p_1 holding price p_2 fixed at p_2^* and finding the corresponding tangency of the price line with slope $-p_1/p_2^*$ to the production possibility curve. This slope is the social marginal cost of q_1 in terms of q_2, that is, the value of q_2 that must be given up to gain an additional unit of q_1 at prices p_1 and p_2^*. As the amount of q_1 increases, social marginal cost increases. The demand curve in Figure 3.6(b) is the graph of the absolute value of the slope of the Scitovsky indifference curve C in Figure 3.6(a). It corresponds to varying the price p_1 holding price p_2 fixed at p_2^* and finding the corresponding tangency of the price line with slope $-p_1/p_2^*$ to the Scitovsky curve. Thus, the Scitovsky curve has a social marginal willingness-to-pay (WTP) interpretation. That is, at each point on curve C the slope is the maximum amount of q_2 society is willing to give up to gain an additional unit of q_1 at prices p_1 and p_2^*. As society has more of q_1, the social marginal WTP, or social marginal benefit (MB), declines.

At the Pareto optimal level of q_1, denoted by q_1^* in Figure 3.6, the marginal WTP is just equal to the marginal cost of q_1, so it is impossible to identify any potential economic social gains in moving from this point. In Figure 3.6(b) in particular, this result is noted by considering movements to the right and left of q_1^*. For example, for a movement to the right, say to q_1^1, the marginal cost is greater than marginal WTP and, hence, losses are associated with moving from q_1^* to q_1^1. To the left, say at q_1^2, marginal WTP exceeds marginal cost so net social benefits are possible in moving from q_1^2 to q_1^*. Finally, note that under the assumptions of Section 2.5 the competitive mechanism results in a market equilibrium at product price p_1^* and quantity q_1^*, which attains Pareto efficiency given price p_2^* for good q_2.

Chapters 4 through 7 use the approach of Figure 3.6(b) assuming the prices in other

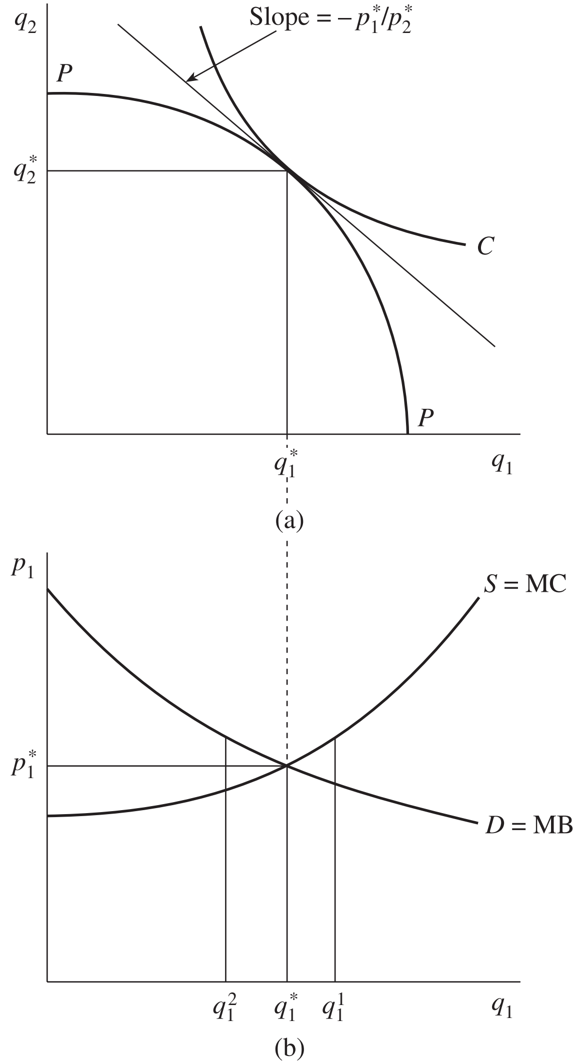

Figure 3.6

markets are fixed. Focusing on a single market while assuming that equilibrium in other markets is unaffected constitutes a partial equilibrium approach in contrast to the framework of Chapters 2 and 3, which uses a general equilibrium approach. A partial equilibrium approach is useful for illuminating how specific policies affect specific markets and groups of consumers and producers in specific markets. However, one must bear in mind that specific policies with specific effects in a given market can have additional general equilibrium implications in other markets. These are considered in Chapter 9.[10]

10. The results illustrated in the simple graphical model of Figure 3.6 are developed rigorously in the Appendix to Chapter 9 in the context of a market economy with many markets.

3.7 SUMMARY

This chapter has focused on the compensation principle and the social welfare function as devices to aid policy-makers in using resources optimally. The compensation principle states that state *B* is preferred to state *A* if, in making the move from state *A* to state *B*, the gainers can compensate the losers such that everyone could be made better off – that is, if the WTP of the gainers exceeds the WTA of the losers. The principle is based on potential rather than actual compensation. Some could actually be made worse off from a policy change, yet the change would be supported if the gainers *could* have compensated the losers so that everyone could have been better off. Because the principle is based on potential rather than actual gains, two problems arise: the reversal paradox and the intransitivity problem. However, even though the criterion is based on potential gains, these problems can arise only if no first-best bundle is considered.

The concept of a utility possibility curve was introduced, and a parallel was drawn between the utility possibilities approach to welfare economics versus that based on production possibility frontiers and Scitovsky indifference curves. The notion of a potential welfare criterion was introduced. If this criterion is adhered to by policy-makers, all possible bundles of goods have to be considered together with all possible distributions of these bundles. Such an approach is usually not empirically practical (although it is considered further in Section 8.3).

Because the compensation principle cannot rank first-best bundles, the concept of a welfare function was introduced. If such a function were available and agreed upon, the optimum organization could be obtained. But because agreement on such a function cannot be reached, the compensation principle is apparently the most widely applicable, yet also empirically practical, criterion. However, one of the problems with the principle is that it is based on potential rather than actual gains. Thus, in any policy context, the payment of compensation is a matter that must be decided by policy-makers endowed with the authority to determine income distributional issues.