Chapter 15
An Exchange Economy

15.1 Introduction
What economists call a pure exchange economy, or more simply an exchange economy, is a model of an economy with no production. Goods have already been produced, found, inherited, or endowed, and the only issue is how they should be distributed and consumed. Even though this model abstracts from production decisions, it illustrates important questions about the efficiency or inefficiency of allocations of goods among consumers, and provides important answers to those questions.

In this chapter, we start with a very simple model of an exchange economy, and we will discuss Pareto optimality or Pareto efficiency for allocations of goods among consumers. Then we turn to the role of markets, and discuss market or competitive equilibrium allocations. Finally, we discuss the extremely important connections between markets and efficiency in an exchange economy. These connections between markets and efficiency are among the most important results in economic theory, and are appropriately called the fundamental theorems of welfare economics.

15.2 An Economy with Two Consumers and Two Goods
We will study the simplest possible exchange economy model, with only two consumers and two goods. A model of exchange cannot get much simpler, because if there were only one good or one person, there would be no reason for trade. However, even though our model is extremely simple, it captures all the important issues, and it generalizes easily.

Let us suppose that there are two consumers. In recognition of Daniel Defoe’s early novel Robinson Crusoe (published in 1719), we call them Robinson and Friday. We abbreviate Robinson R; for Friday we use F. We assume there are only two consumption goods on their island, bread (good x) and rum (good y). In
Chapter 15. An Exchange Economy

This model of exchange, we are abstracting from the fact that Robinson and Friday produce rum (or somehow have acquired a stock of it), and produce bread (by making flour, mixing, and baking). Therefore we assume that there are fixed totals of rum and bread available, and that the only issue is how to distribute those totals among the two consumers.

Here’s how the distribution of the two goods works. The two consumers start with initial endowments of the goods; then they make trades. We let \( X \) represent the total quantity of good \( x \), bread, that is available. We let \( Y \) represent the total quantity of good \( y \), rum, that is available. In general, if we are talking about an arbitrary bundle of goods for Robinson, we show it as \((x_R, y_R)\), where \( x_R \) is his quantity of bread and \( y_R \) is his quantity of rum. An arbitrary bundle of goods for Friday is \((x_F, y_F)\).

Robinson has initial endowments of the two goods, as does Friday. We will use the “naught” superscript (that is, \( ^0 \)) to indicate an initial quantity. Robinson’s initial bundle of goods is \((x_R^0, y_R^0)\). Friday’s initial bundle of goods is \((x_F^0, y_F^0)\). The quantities of the two goods in the initial bundles must be consistent with the assumed totals of bread and rum. That is,

\[
X = x_R^0 + x_F^0 \quad \text{and} \quad Y = y_R^0 + y_F^0.
\]

Moreover, if they start with their initial quantities and then trade, any bundles they end up with must also be consistent with the given totals. That is, if they end up at \(((x_R, y_R), (x_F, y_F))\), it must be the case that

\[
X = x_R + x_F \quad \text{and} \quad Y = y_R + y_F.
\]

Robinson’s preferences for bread and rum are represented by the utility function \( u_R(x_R, y_R) \), and, similarly, Friday’s preferences are represented by \( u_F(x_F, y_F) \). That is, we assume that each consumer’s utility depends only on his own consumption bundle. The utility functions \( u_R \) and \( u_F \) will generally be different, and unrelated to the initial bundles that Robinson and Friday happen to have. The facts that preferences are generally different, and initial bundles are also generally different, make mutually beneficial trade probable.

To show our simple exchange economy with a graph, we use a diagram first suggested (in 1881) by the great Anglo-Irish economist Francis Ysidro Edgeworth (1845–1926). (Actually, Edgeworth didn’t really invent this diagram; the version we use today is due to the English economist Arthur Bowley (1869–1957).) This graph is called an Edgeworth box diagram. We show it in Figure 15.1. In the figure, the initial endowment is given by the point \( W \). That is, \( W \) is the allocation of bread \( x \) and rum \( y \) giving Robinson the bundle \((x_R^0, y_R^0)\) and giving Friday the bundle \((x_F^0, y_F^0)\). There are two indifference curves shown in the figure: \( I_R \) belongs to Robinson, and \( I_F \) belongs to Friday. The small arrows attached to those indifference curves indicate the directions of increasing utility.

The novel feature of the Edgeworth box diagram, which makes it different from other diagrams we have used, is that it has two origins. The lower left origin is
15.3 Pareto Efficiency

In previous chapters, when we analyzed the welfare properties of competitive markets, monopoly, and duopoly, we looked at efficiency in terms of consumers' and producers' surplus. The sum of these two surpluses, or the social surplus,
represents the total net benefit created in a market, to buyers and sellers, measured in money units such as dollars. Social surplus is a measure of the "size of the economic pie," created by the production and trade of some good. A market for a good is inefficient if there is a way to make that pie bigger (e.g., the standard monopoly case), and it is efficient if there is no way to make it bigger (e.g., the standard competitive case).

All this assumed that consumers' surplus is well defined, which in turn required some special assumptions about preferences. Note that this kind of analysis focused on one good under study, and ignored what might have been happening in other markets for other goods, for labor and savings, and so on. It was therefore what is called partial equilibrium analysis; models that study one good in this fashion are called partial equilibrium models.

We are now, however, looking at a simple model of exchange, without production. If we measure the size of the economic pie in terms of total quantities of bread and/or rum, the size cannot change because these total quantities are fixed. There is no money in the model (at least not yet), so it would not be easy to measure the size of the pie in money units. We might try to measure the economic pie in utility units, but we know that it would probably be wrong to try to add together Robinson's utility and Friday's utility. How, then, can we decide when an allocation of the fixed quantities of bread and rum between the two consumers is efficient (or when it is not)?

The solution to this problem was developed by the Italian economist Vilfredo Pareto (1848-1923), so we call the central concept Pareto optimality or Pareto efficiency. Here are some important definitions.

First, we need to be careful about which allocations of bread and rum are possible and which are not. We will say that a pair of bundles of goods, \((x_R, y_R)\) and \((x_F, y_F)\), is a feasible allocation if all the quantities are nonnegative and if

\[
X = x_R + x_F \quad \text{and} \quad Y = y_R + y_F.
\]

That is, a feasible allocation is one in which the goods going to Robinson and Friday add up to the given totals. In fact, the feasible allocations in the exchange model are simply the points in the Edgeworth box diagram, no more and no less.

Second, if \(A\) and \(B\) are two feasible allocations, we will say that \(A\) Pareto dominates \(B\) if both Robinson and Friday like \(A\) at least as well as \(B\), and at least one likes it better. If \(A\) Pareto dominates \(B\), we call a move by Robinson and Friday from \(B\) to \(A\) a Pareto move.

Third, a feasible allocation is not Pareto optimal if there is a different feasible allocation that both consumers like at least as well, and that is preferred by at least one of them. That is, a feasible allocation is not Pareto optimal if there is a Pareto move from it. Note that any Pareto move would get a unanimous vote of approval (possibly with an abstention).

Fourth, and finally, a feasible allocation is Pareto optimal or Pareto efficient if there is no feasible allocation that both consumers like at least as well, and
15.3 Pareto Efficiency

that is preferred by at least one of them. That is, a feasible allocation is Pareto optimal if there is no Pareto move from it. The reader can see a nonoptimal feasible allocation in Figure 15.1 — the initial allocation \( W \) — and there are many points in the Edgeworth box diagram (the lens-shaped area to the northwest of \( W \)) that Pareto dominate \( W \). (Of course \( W \) is not the only nonoptimal allocation in Figure 15.1.)

These Pareto-related definitions can easily be extended to exchange economies with any number of consumers and any number of goods, and can be extended, less easily, to any kind of economic model, including models with production as well as exchange. These definitions are not restricted to models of markets with just one good ("partial equilibrium models"). They are very general, and are useful in general equilibrium models, that is, models that consider supply and demand in all markets simultaneously.

Our pure exchange model, with two people and two goods, is a very simple example of a general equilibrium model.

Now think about a feasible allocation that is not Pareto optimal. It is obviously undesirable for society to be at that allocation, as there are other feasible allocations that are unambiguously better, in the sense that a move from the given nonoptimal allocation to the alternative would get unanimous consent. Obviously, if an economy is at a non-Pareto optimal point, it should move to something better. None, however, that Pareto optimality has nothing to do with considerations of distributional fairness or equity. That is, a non-Pareto optimal allocation may be much more equal than a Pareto optimal one. In fact, allocating all the bread and rum to Robinson (for example) is Pareto optimal, as there is no Pareto move away from that totally lopsided and unfair allocation. Moreover, giving both Robinson and Friday exactly half the bread and half the rum, the allocation that is the most equal of all the feasible allocations, is probably not Pareto optimal.

With all this said, we now return to our Edgeworth box diagrams. They should make the mysteries of Pareto optimality and non-optimality clear.

Feasible Allocations and the Edgeworth Box

Suppose we have a bundle of goods \((x_R, y_R)\) for Robinson and another bundle \((x_F, y_F)\) for Friday. For the pair of bundles \(((x_R, y_R), (x_F, y_F))\) to be a feasible allocation, the numbers must add up to the totals \(X\) and \(Y\). Consider Figure 15.2. The two bundles are shown in it, but the quantities of bread (on the horizontal axis) and rum (on the vertical axis) do not add up to the total quantity of bread available, \(X\) (the width of the box), or the total quantity of rum available, \(Y\) (the height of the box). In other words, if you are interested in finding Pareto optimal allocations of bread and rum, don’t even think about the pair of bundles \((x_R, y_R)\) and \((x_F, y_F)\) shown in Figure 15.2, because that pair of bundles is just not possible.
In Figure 15.3, we show another pair of bundles of goods whose totals do not add up to X and Y. However, this time the totals fall short. We still consider this pair of bundles \(((x_R, y_R), (x_F, y_F))\) nonfeasible and therefore non-Pareto optimal. (Some economists would pronounce \(((x_R, y_R), (x_F, y_F))\) “feasible,” because one could throw away some bread and some rum (the excesses of each good), starting at X and Y, and get to the lesser totals. But even if one takes this approach, \(((x_R, y_R), (x_F, y_F))\) still would not be Pareto optimal, because it is Pareto dominated by another allocation formed by starting with \(((x_R, y_R), (x_F, y_F))\), and then adding back half the excesses to each consumer’s bundle.)

The moral of Figures 15.2 and 15.3 is that in the exchange economy model, an allocation of bread and rum must be feasible before one can decide whether it is Pareto optimal. To be feasible, it must be the case that

\[ X = x_R + x_F \quad \text{and} \quad Y = y_R + y_F. \]

Tangencies of Indifference Curves and the Edgeworth Box

From this point on, we consider only feasible allocations. That is, we look only at pairs of bundles \(((x_R, y_R), (x_F, y_F))\) that can be represented by single points in the Edgeworth box diagram. The point W in Figure 15.1 was such a feasible allocation. We will draw a fresh figure, Figure 15.4, with a similar point, labeled P. In the figure, the indifference curves \(I_R\) and \(I_F\) cross at the point \(P = ((x_R, y_R), (x_F, y_F))\). The arrows on the indifference curves show the directions of increasing utility. The point \(P\) cannot be Pareto optimal, because a move to the interior of the lens-shaped area would make both consumers better off. (Moving to the other end of the lens,
where $I_R$ and $I_F$ cross again, would leave both of them exactly as well off; moving to the edges of the lens would make one person better off and the other exactly as well off as at the highlighted point.)

We now have a tentative conclusion. If the indifference curves of the two consumers cross at a point in an Edgeworth box diagram, that point must be non-Pareto optimal, and there must be other points, other feasible allocations, that Pareto dominate it. Note, by the way, that if two indifference curves actually cross at a point in the box (rather than just touch each other), then that point must be in the interior of the box. That is, it must be the case that $x_R$, $y_R$, $x_F$, and $y_F > 0$.

If Robinson and Friday are at a point such as $P$ in Figure 15.4, they can make trades that benefit one or both, and harm neither. That is, they can make Pareto moves. If they are free to trade, aware of the feasible allocations, and in touch with their preferences or utility functions, they will probably continue to trade until they can no longer make Pareto moves; that is, they will trade to a Pareto optimal allocation. Moreover, if the point at which they end up is in the interior of the Edgeworth box diagram, it cannot be a point at which Robinson's and Friday's indifference curves cross. Rather, assuming that the indifference curves are smooth and do not have kinks, the point at which they end up, when further Pareto moves are impossible, must be a tangency point.

In short, in the interior of the Edgeworth box diagram, the Pareto optimal points must be points of tangency between the indifference curves of Robinson and Friday. Figure 15.5 shows one such Pareto optimal point, identified as $Q = ((x_R, y_R), (x_F, y_F))$. Four arrows are drawn from $Q$. The one pointing

![Figure 15.3](image-url)
Chapter 15. An Exchange Economy

Fig. 15.4. A point in the Edgeworth box diagram that is not Pareto optimal.

Fig. 15.5. The allocation $Q$ is Pareto optimal.

The northeast suggests a move that would make Robinson better off, but would make Friday worse off. The one pointing southwest suggests a move that would make Friday better off, but would make Robinson worse off. A move in the direction of each of the other arrows would make both consumers worse off. Therefore, there is no Pareto move away from $Q$, which means that $Q$ must be Pareto optimal.

All this leads to a necessary condition for Pareto optimality for points in the interior of the Edgeworth box diagram. For such a point to be Pareto optimal, the
slopes of the indifference curves of Robinson and Friday must be equal at the point. That is, Robinson's marginal rate of substitution of good $y$ for good $x$ must equal Friday's marginal rate of substitution of $y$ for $x$. This gives

$$MRS^R = MRS^F,$$

which in turn gives

$$\frac{MU_x^R}{MU_y^R} = \frac{MU_x^F}{MU_y^F}.$$

Note that the $R$ superscript is for Robinson and the $F$ superscript is for Friday.

The set of Pareto optimal allocations in the Edgeworth box diagram is called the contract curve. The name is fitting, for these are the allocations that could potentially be outcomes of trading contracts. That is, Robinson and Friday would be likely to agree to a contract that would take them from the initial allocation $W$ to the contract curve. Of course, where they end up on the contract curve depends on the location of the starting allocation $W$, and it may also depend on their bargaining abilities. If $W$ gives most of the bread and most of the rum to Robinson, they will end up somewhere on the contract curve where Robinson still has most of the bread and most of the rum. In Figure 15.6, we show a contract curve. In the interior of the Edgeworth box diagram it is the set of tangency points. We show an initial allocation $W$ that gives most of the bread to Robinson and most of the rum to Friday. In this exchange economy, Robinson and Friday will trade to the contract curve, but not to anywhere on the contract curve. They will want to make Pareto moves. This means that neither should end up worse off than they were at the initial
allocation $W$. The part of the contract curve where neither is worse off than they were at $W$ is between allocations $A$ and $B$. This is called the core.

Most economists believe that Pareto moves are unambiguously good, and that Pareto optimality is desirable, as any non-Pareto optimal point is unambiguously inferior to some Pareto optimal one. Most economists who look at the Edgeworth box diagram agree that it would be a good thing to end up on the contract curve. There is, of course, disagreement about distribution, so we do not claim that the Pareto optimal point $A$ is better than (or worse than) the Pareto optimal point $B$. But we do agree that it would be a good thing to end up at some Pareto optimal point—at some point on the contract curve. We have suggested that in the exchange economy model, our two traders, starting at some initial allocation $W$, can simply trade, or barter, to get to the contract curve. Is there another way to get there? The answer is "Yes," through market trade.

15.4 Competitive or Walrasian Equilibrium

We now model a competitive market in our simple two-person, two-good Robinson/Friday economy. This theory was first developed by the French economist Leon Walras (1834–1910). The market equilibrium idea we describe is called a competitive equilibrium or a Walrasian equilibrium. The connections between market equilibria and Pareto optimality were rigorously analyzed in the 1950s, especially by the American economists Kenneth Arrow (1921–) and Lionel McKenzie (1919–) and the French economist Gerard Debreu (1921–2004).

Here is the story of Walrasian equilibrium. Imagine that an auctioneer lands on the island with Robinson and Friday. The auctioneer has no bread and no rum, nor does he have any desire to consume any. His sole function is to create a market in which people can trade the two goods. He does this by calling out prices for the two goods. He starts by announcing $p_x$, the per-unit bread price, and $p_y$, the per-unit rum price. He announces that he will buy or sell any quantities of bread and/or rum, at those prices. He asks Robinson and Friday, "What do you want to do at those prices?"

In our model of a competitive market economy, we assume that Robinson and Friday take those prices as given and fixed, unaffected by their actions. (This is obviously a little unrealistic when we are talking about just two consumers, but the model is meant to be extended to cases in which there are many consumers, when the assumption of competitive behavior becomes plausible.) Robinson and Friday hear the Walrasian auctioneer announce a pair of prices $(p_x, p_y)$, and they understand that they should tell him what bundle they want to consume, based on those prices.

Our traders have no money in the bank or in their pockets; they have only their initial bundles. Robinson and Friday hear the announced prices and know the bundles they start with. If Robinson starts with 10 loaves of bread, and decides he
wants to consume 12 loaves, he will go to the auctioneer and swap some of his rum for the extra 2 loaves of bread. What exactly is his budget constraint? We could figure it in terms of such a swap; it would then be “value of bread acquired = value of rum given up,” or

\[(x_R - x_R^0)p_x = (y_R - y_R)p_y.\]

With a little rearranging, this gives

\[p_x x_R + p_y y_R = p_x x_R^0 + p_y y_R^0.\]

Alternatively, we could derive Robinson’s budget constraint by realizing that in a world in which consumers do not have money income, what substitutes for income in the budget constraint is the value of the bundle the consumer starts out with. Robinson’s budget constraint should then say “value of his desired consumption bundle = value of the bundle he starts with,” which also gives

\[p_x x_R + p_y y_R = p_x x_R^0 + p_y y_R^0.\]

Recall that the Walrasian auctioneer has called out some prices, and asked Robinson and Friday, “What do you want to do at these prices?” Robinson, of course, wants to maximize his utility, or get to the highest indifference curve, subject to his budget constraint. That is, he wants to maximize

\[u_R(x_R, y_R)\]

subject to the constraint

\[p_x x_R + p_y y_R = p_x x_R^0 + p_y y_R^0.\]

Think of the slope of the budget line implied by this budget constraint. The absolute value of the slope is \(p_x/p_y\), and the budget line must go through the initial bundle \((x_R^0, y_R^0)\). All this leads Robinson to conclude that he wants to consume some bundle; call it \(A_R\) for now. Robinson tells the auctioneer that based on the announced prices, he wants to consume \(A_R\).

Friday goes through the same exercise, and he ends up telling the auctioneer that he wants to consume \(B_F\).

In Figure 15.7 we plot the results. There is one budget line going through the initial allocation \(W\). We do not have two separate lines, one for Robinson and the other for Friday. This is because, first, either trader’s line must go through \(W\), which represents both their initial bundles, and second, because the auctioneer called out only one set of prices, there is only one possible price ratio \(p_x/p_y\) and only one possible slope. In the figure, we show the bundle \(A_R\) that Robinson would like to consume, and the bundle \(B_F\) that Friday would like to consume.

Now it is time for the Walrasian auctioneer to act. He asks himself, “Is it possible for Robinson to consume \(A_R\) and for Friday to consume \(B_F\)?” The reader should immediately see the answer: “No,” because \((A_R, B_F)\) is not a feasible allocation.
Fig. 15.7. At these prices, there is excess supply of good $x$ and excess demand of good $y$. The Walrasian auctioneer should announce new prices with a lower relative price for $x$, $p_x/p_y$.

The totals do not add up to $X$ and $Y$. In particular, the amount of bread that the two want to consume is less than the amount $X$ that is available, and the amount of rum that the two want to consume is greater than the amount $Y$ that is available. Therefore, $(A_R, B_F)$ is just not possible.

The Walrasian auctioneer sees this. He says to himself, “The $(p_x, p_y)$ I announced must be changed. There is excess supply of bread and excess demand for rum. I must lower the relative price of bread $p_x/p_y$.” He then tells Robinson and Friday that there will be no trading at the previously announced prices. Instead, he announces a new pair of prices, for which $p_x/p_y$ is a little lower than the first pair of prices. (For instance, if the original $(p_x, p_y)$ was (2, 1), he announces new prices (1.75, 1).) He tells Robinson and Friday to forget about the bundles they wanted to consume at the previous pair of prices. Instead, they should now tell him what bundles they want to consume at his newly announced prices. Robinson and Friday then figure out what bundles they want to consume at the new prices, and duly report back to the auctioneer.

The auctioneer, Robinson, and Friday continue this price-to-desired-consumption-bundles-to-price process until, finally, they end up with a pair of prices and desired consumption bundles that work. That is, the process continues until Robinson and Friday tell the Walrasian auctioneer that based on his latest price combination $(p_x^*, p_y^*)$, they want to consume certain bundles $A_R^*$ and $B_F^*$, and those bundles are consistent with the given totals of bread and rum; they are
a feasible allocation, a single point in the Edgeworth box diagram. That is, at the pair of bundles $A^*_R$ and $B^*_F$, for each of the goods, total demand = total supply.

Once this end has been reached, the Walrasian auctioneer makes his final announcement to Robinson and Friday: "We're finally there. Make the trades at the $(p^*_R, p^*_F)$ prices, either through me as an intermediary or directly between yourselves. Then consume and enjoy!"

(We have been somewhat casual about the nature of the dynamic price adjustment process. Analysis of convergence for the process is beyond our scope.)

The process we described above is called a Walrasian auctioneering process or Walrasian process or tatonnement process. (The word "tatonnement" is French for "groping.") The end result is called a competitive equilibrium or Walrasian equilibrium. The pair of prices at which it ends up, $p^*_R$ and $p^*_F$, are called the competitive equilibrium prices. The Walrasian process produces the equilibrium prices and a pair of consumption bundles $A^*_R$ and $B^*_F$, such that $A^*_R$ maximizes Robinson's utility subject to his budget constraint with the equilibrium prices, $B^*_F$ maximizes Friday's utility subject to his budget constraint with the equilibrium prices, and $(A^*_R, B^*_F)$ is a feasible allocation; that is, the desired total consumption of each good equals the total supply of that good. The allocation $(A^*_R, B^*_F)$ is called a competitive equilibrium allocation.

Figure 15.8 shows a competitive equilibrium. Note the crucial difference between Figure 15.7 and Figure 15.8: In Figure 15.7 the desired consumption bundles are two distinct points in the Edgeworth box, which means they are not a feasible allocation; there is excess supply of bread and excess demand for rum. This suggests the relative price of bread $p_x/p_y$, should fall. That is, the budget line should get flatter. Figure 15.8 has a flatter budget line, and in that figure the desired consumption bundles do coincide in the Edgeworth box. They constitute a feasible allocation. Supply equals demand for each good.

Note two extremely important facts about $(A^*_R, B^*_F)$ in Figure 15.8. First, the competitive equilibrium allocation is a tangency point for the two indifference curves shown. That means it is on the contract curve. It is Pareto optimal. Second, a look at Figures 15.6 and 15.8 together should convince you that the competitive equilibrium allocation is in the core.

15.5 The Two Fundamental Theorems of Welfare Economics

The relationships between free markets and efficiency, and between market incentives and national wealth, have been written about since the time of Adam Smith (1723–1790), who published *The Wealth of Nations* in 1776. Smith's arguments were neither formal nor mathematical; the formal and mathematical analysis was developed in the late nineteenth and mid twentieth centuries. We now call the two basic results that relate Pareto optimality and competitive markets the first and second fundamental theorems of welfare economics. Figure 15.8 illustrates
Chapter 15. An Exchange Economy

The first fundamental theorem in our simple pure exchange model, with only two people and two goods. The figure shows that a competitive equilibrium allocation is Pareto optimal. That result easily extends to exchange models with any number of people and any number of goods, as well as to economic models with production as well as exchange. The result requires only a few assumptions; in particular, we must assume that there are markets and market prices for all the goods, that all the agents are competitive price takers, and that any individual’s utility depends only on his or her own consumption bundle, and not on the consumption bundles of other individuals. (Similarly, if there are firms, we must assume that they are all competitive price takers, and that any firm’s production function depends only on that firm’s inputs and outputs.) We now state the first fundamental theorem, for a general exchange economy.

First Fundamental Theorem of Welfare Economics

Suppose that there are markets and market prices for all the goods, that all the people are competitive price takers, and that each person’s utility depends only on her own bundle of goods. Then any competitive equilibrium allocation is Pareto optimal. In fact, any competitive equilibrium allocation is in the core.

This is an extremely important result, because it suggests that a society that relies on competitive markets will achieve Pareto optimality. It is a remarkable result because, although there are many competitive allocations, most allocations in an exchange model are actually not Pareto optimal. The reader should look back at Figure 15.6 and think about throwing a dart at that Edgeworth box diagram,
hoping to hit the contract curve. What are the odds you will hit it? At least in theory the odds are zero, because a line has zero area. Therefore, ending up at a Pareto optimal allocation is not easy, and the fact that the market mechanism does it is impressive. Moreover, the market mechanism is cheap (it requires only a Walrasian auctioneer, in theory, or perhaps something such as eBay, in reality). It does not require that some central power learn everybody’s utility function (which would be terribly intrusive and dangerous) and then make distributional decisions; it requires only publicly known prices that move in response to excess supply or demand. In short, the competitive market mechanism is relatively cheap, relatively unobtrusive, relatively benign, and remarkably effective. This is what the first fundamental theorem of welfare economics helps us understand.

However, one important shortcoming of the first fundamental theorem is that the location of the competitive equilibrium allocation is highly dependent on the location of the initial allocation. In other words, if we start at an initial allocation that gives Robinson most of the bread and rum, we will end up at a competitive equilibrium allocation that gives Robinson most of the bread and rum. Or, more generally, if a society has a very unequal distribution of talents and abilities and initial quantities of various goods, it will end up with a competitive equilibrium that, while Pareto optimal, is very unequal. What can be done? This is where the second fundamental theorem comes in.

The second fundamental theorem of welfare economics uses all the assumptions of the first theorem, and adds an additional one, convexity. In particular, at least for the exchange version of the second fundamental theorem, we will assume that the traders have convex indifference curves. (This is, in fact, how we drew the indifference curves in Figures 15.4 through 15.8.) Here is what the theorem says. Suppose the initial allocation in society is very skewed, very unfair, and therefore a competitive equilibrium based on it would be very unfair. Suppose that people in society have decided that there is a different, perhaps much fairer Pareto optimal allocation, that they want to get to. However, they want to use the market mechanism mostly to get to that desired Pareto optimal point; they do not want a dictator announcing what bundle of goods each and every person should consume. Is there a slightly modified market mechanism that will get society from the initial allocation to the target Pareto optimal one? The answer is “Yes.”

Figure 15.9 illustrates the theorem. The initial allocation $W$ gives all the bread and most of the rum to Robinson. Given this initial allocation, the Walrasian mechanism we described earlier, if left alone, will produce a competitive equilibrium that gives most of the bread and most of the rum to Robinson. This allocation is labeled *Laissez-faire* in the figure. This is the outcome of the unshackled free market. ("Laissez-faire" is French for "let do"; that is, let the market do its thing.) But it is very unfair; it leaves Friday a pauper. A more equitable goal would be the Pareto optimal allocation labeled *Target*. Can the market, with a relatively small fix, be used to get society to Target?
Referring now to Figure 15.9, this is how the market mechanism is modified to get the economy to Target, the Pareto optimal allocation that is more equitable than what the market would produce by itself. Because Target is Pareto optimal, and is in the interior of the Edgeworth box diagram, it must be a tangency point of the two traders’ indifference curves. (The argument is more complicated if Target is not an interior point.) Because it is a tangency point, we can draw a tangent line like the one in Figure 15.9. From the slope of the tangent line, we can figure out what price ratio \( p_x^*/p_y^* \) we are going to need. One of the prices could be arbitrarily set to 1, and then the required price ratio would give the other price. (A good whose price is set equal to 1 is called a *numeraire* good.) This gives the required pair of prices \((p_x^*, p_y^*)\).

Now, the government must step in and introduce some *lump-sum taxes and transfers*. These are imposed on Robinson and Friday. They are “lump sum” because they are independent of the quantities of goods the parties want to consume. A tax (money taken from a person) will be represented by a negative number; a transfer (or subsidy) will be represented by a positive number. Let \( T_R \) be Robinson’s tax or transfer, and let \( T_F \) be Friday’s. The government is not creating or destroying wealth, so we require that

\[
T_R + T_F = 0.
\]

Whatever the government taxes away from one party will be quickly sent to the other party. The government now sends Robinson a message: “Put \( T_R \) onto the right-hand side of your budget constraint, and assume prices \((p_x^*, p_y^*)\) for the two
goods. If $T_R$ is a negative number, too bad for you. You’ve lost some money; send it to us. If it’s a positive number, good for you. You’ve gained some money; we’ll be wiring it to you today.” The government sends Friday a similar message about $T_F$. The budget constraints of Robinson and Friday now become

$$p_x^*x_R + p_y^*y_R = p_x^0x_R + p_y^0y_R + T_R$$

and

$$p_x^*x_F + p_y^*y_F = p_x^0x_F + p_y^0y_F + T_F.$$ 

Robinson and Friday now choose their desired consumption bundles, based on these budget constraints. If the government sets the taxes and transfers at the right level, Robinson and Friday will end up at the desired Pareto optimal point, the point Target.

In short, by properly setting lump-sum taxes and transfers, society can get from any initial allocation, no matter how inequitable, to a more desirable Pareto optimal allocation without abandoning the use of the market mechanism. The second fundamental theorem of welfare economics says all this is possible. Here is a more formal statement.

**Second Fundamental Theorem of Welfare Economics**

Suppose that there are markets for all the goods, that all the people are competitive price takers, and that each person’s utility depends only on his or her own bundle of goods. Suppose further that the traders have convex indifference curves. Let Target be any Pareto optimal allocation.

Then there are competitive equilibrium prices for the goods, and a list of lump-sum taxes and transfers for the people, which sum to zero, such that when the budget constraints based on these prices are modified with these taxes and transfers, Target is the resulting competitive equilibrium allocation.

Loosely speaking, the first fundamental theorem of welfare economics says that any competitive equilibrium is Pareto optimal, and the second says that any Pareto optimal point is a competitive equilibrium, given the appropriate modification of the traders’ budget constraints. The second theorem needs an additional assumption (convexity), and relies heavily on the budget constraint modifications. However, the existence of the second theorem allows all economists to agree, more or less, “We like the market mechanism; it gets us Pareto optimality.” Conservative economists tend to say, “The market’s great, don’t touch it; let’s go to the Pareto optimal outcome it gives us.” Liberal economists tend to say, “The market’s great, but the initial allocation is terrible; let’s use some taxes and transfers to fix the inequities, and then let’s go to the Pareto optimal outcome it gives us.”

This debate is one of the things that makes life interesting for economists, and for many others.
15.6 A Solved Problem

The Problem
Consider a pure exchange economy with two consumers, 1 and 2, and two goods, x and y. Consumer 1’s initial endowment is \( w_1 = (1, 0) \); that is, 1 unit of good x and 0 units of good y. Consumer 2’s initial endowment is \( w_2 = (0, 1) \); that is, 0 units of good x and 1 unit of good y. Let \( W = (w_1, w_2) \) represent the initial allocation. Consumer \( i \)’s utility function (for \( i = 1, 2 \)) is \( u_i(x_i, y_i) = x_i y_i \), where \((x_i, y_i)\) represents \( i \)’s consumption bundle.

(a) Show this economy (with some indifference curves and the initial endowments) in an Edgeworth box diagram.

(b) Write down the equations that describe the Pareto efficient allocations. Identify them in the Edgeworth box. Is the initial endowment point Pareto efficient? Why or why not?

(c) Calculate the competitive equilibrium of this pure exchange economy. You should indicate final consumption bundles for each agent, and the equilibrium prices \( p_x^* \) and \( p_y^* \). (Remember that you can normalize the price of one good to be 1.)

(d) For any pair of prices \((p_x, p_y)\), consumers 1 and 2 can figure their desired consumption levels and the net amounts of good x and good y that they want to buy or sell. For example, consumer 1’s net demand for good x will be \( x_1 - 1 \). (This is a negative number, meaning that he will want to sell some of his initial 1 unit of x.) Similarly, consumer 2’s net demand for good x will be \( x_2 - 0 \). Adding over both consumers gives the total net demand for good x, or the excess demand for x measured in units of x. (This might be positive or negative. If it is negative, there is excess supply of good x.) Multiplying by \( p_x \) would give the excess demand for x measured in dollars.

Assume that \((p_x, p_y)\) is any pair of positive prices. (Note that this is any pair of prices, not just the competitive equilibrium prices.) Show that the sum of excess demand for good x in dollars and excess demand for good y in dollars must be zero. This kind of result was first formally established by Leon Walras, and is therefore called Walras’ law. Walras’ law can be put this way: the sum of market excess demands, over all markets, measured in currency, must be zero.

The Solution
(a) We will not draw the Edgeworth box diagram; it is very similar to Figure 15.8. However, we will describe the diagram. Please sketch the diagram as you read this, if you have not already drawn it. The Edgeworth box diagram is a square, one unit on each side. Consumer 1’s origin is the lower left-hand corner; consumer 2’s origin is the upper right-hand corner. The initial point \( W \) is the lower right-hand corner of the box. Indifference curves are generally symmetric hyperbolas, symmetric around the diagonal of the box that goes from consumer 1’s origin to consumer 2’s origin, that is, from lower left to upper right. However, the indifference curves that go through the initial point \( W \) are “degenerate” hyperbolas; this means that for consumer 1, for instance, the indifference curve through his initial bundle \((1, 0)\) is given by \( x_1 y_1 = 0 \); graphically, this is his horizontal axis plus his vertical axis.
(b) The Pareto optimal points in this example are points of tangency between indifference curves of the two consumers. Tangency requires that consumer 1’s marginal rate of substitution equal consumer 2’s marginal rate of substitution. Consumer i’s marginal rate of substitution is

\[ MRS^i = \frac{MU^i_x}{MU^i_y} = \frac{y_i}{x_i}. \]

Setting the two consumers’ marginal rates of substitution equal, and then substituting \( 1 - x_1 \) for \( x_2 \) and \( 1 - y_1 \) for \( y_2 \), gives

\[ \frac{y_1}{x_1} = \frac{y_2}{x_2} = \frac{1 - y_1}{1 - x_1}. \]

This leads directly to

\[ x_1 = y_1. \]

Therefore, the set of Pareto optimal points, that is, the contract curve, is simply the upward-sloping diagonal of the box diagram, from consumer 1’s origin to consumer 2’s origin. The initial point \( W \) is obviously not efficient; it is not on the contract curve. In fact, any move from \( W \) into the interior of the Edgeworth box diagram would make both consumers better off.

(c) At a competitive equilibrium in the interior of an Edgeworth box diagram, the price ratio \( p_x^* / p_y^* \), consumer 1’s marginal rate of substitution, and consumer 2’s marginal rate of substitution must all be equal. On the contract curve, where \( MRS^1 = MRS^2 \), we found that \( x_1 = y_1 \) must hold. Because \( MRS^1 = y_1/x_1, MRS^1 = 1/1 = 1 \) on the contract curve. Therefore, \( p_x^* / p_y^* = 1 \) at the competitive equilibrium. We are free to set the price for one of the goods (the numeraire good) equal to 1. Let us make good \( y \) the numeraire good. Then \( p_y^* = 1 \), and because \( p_x^* / p_y^* = 1, p_x^* = 1 \) also.

To find the exact location of the competitive equilibrium allocation, we note that the competitive equilibrium budget line must have slope \( p_y^* / p_x^* = 1 \) in absolute value, and must start at the initial allocation \( W \), which is the lower right-hand corner of the box. Therefore, the competitive equilibrium budget line is the diagonal of the box going from the lower right corner to the upper left corner. The competitive equilibrium will be at the intersection of the competitive equilibrium budget line and the contract curve, which we already noted was the lower left to upper right diagonal. It follows that the competitive equilibrium allocation is the exact center of the box.

Consumer 1’s competitive equilibrium bundle is therefore \((1/2, 1/2)\). Consumer 2’s competitive equilibrium bundle is also \((1/2, 1/2)\).

(d) Let \((p_x, p_y)\) be any pair of positive prices, and let \((x_1, y_1)\) and \((x_2, y_2)\) be the corresponding desired consumption bundles of the two consumers. (The assumption of positive prices guarantees that no one wants to consume an infinite amount of \( x \) or \( y \)). We will let \( $ED(x) \) represent the excess demand for \( x \), measured in dollars, and similarly \( $ED(y) \) will represent excess demand for \( y \), measured in dollars. Note that
\[ ED(x) = px(x_1 - 1) + px(x_2 - 0) \]

The sum of excess demands, for goods \( x \) and \( y \) is

\[ ED(x) + ED(y) = px(x_1 - 1) + px(x_2 - 0) + py(y_1 - 0) + py(y_2 - 1) = (pxx_1 + pyy_1 - px) + (pxx_2 + pyy_2 - py). \]

But consumer 1's budget constraint says

\[ qx_1 + pyy_1 = px \times 1 + py \times 0 = px, \]

so the terms in the first set of parentheses sum to zero. Similarly, by consumer 2's budget constraint, the terms in the second set of parentheses sum to zero. Therefore,

\[ ED(x) + ED(y) = 0, \]

which is Walras' law.

**Exercises**

1. There are two goods in the world, tiramisu \((x)\) and espresso \((y)\). Michael and Angelo both consider tiramisu and espresso to be complements; each will consume a slice of tiramisu only if it is accompanied with a cup of espresso, and vice versa. Michael has five slices of tiramisu and a cup of espresso. Angelo has a slice of tiramisu and five cups of espresso.

   (a) Draw an Edgeworth box for this exchange economy. Label it carefully. Mark the original endowment point \( W \).

   (b) Draw Michael's and Angelo's indifference curves passing through the endowment point.

   (c) Can you suggest a Pareto improvement over the original endowment? Mark the new allocation \( W' \). How many slices of tiramisu and cups of espresso will each of them consume at the new allocation?

2. Ginger has a pound of sausages \((x_g = 1)\) and no potatoes \((y_g = 0)\), and Fred has a pound of potatoes \((y_f = 1)\) and no sausages \((x_f = 0)\). Assume Ginger has the utility function \( u_g = x_g^{\alpha} y_g^{1-\alpha} \), Fred has the same utility function, \( u_f = x_f^{\alpha} y_f^{1-\alpha} \), and the parameter \( 0 < \alpha < 1 \) is the same for Ginger and Fred. (You may remember that these are called Cobb-Douglas utility functions.)

   (a) Show that the contract curve is the diagonal of the Edgeworth box.

   (b) Show that at the competitive equilibrium,

   \[ \frac{p_x}{p_y} = \frac{\alpha}{1 - \alpha}. \]

3. Consider an exchange economy with two goods, \( x \) and \( y \), and two consumers, Rin and Tin. Rin's utility function is \( u_r = x, y \), and his endowment is \( \omega_r = (2, 2) \). Tin's utility function is \( u_t = x_t y_t^2 \) and his endowment is \( \omega_t = (3, 3) \). Duncan suggests that there might be a competitive equilibrium at \((x'_r, y'_r) = (4, 1)\), \((x'_t, y'_t) = (1, 4)\), with prices \( p_x = p_y = 1 \).

   (a) Does Duncan's suggested equilibrium allocation have the right totals of the two goods \( x \) and \( y \)? Explain.
(b) Is Duncan’s suggested equilibrium allocation a Pareto improvement over the endowment? Explain.

(c) Write down Rin’s budget constraint given these prices. Solve for Rin’s optimal consumption bundle, \((x^*_r, y^*_r)\).

(d) Write down Tin’s budget constraint given these prices. Solve for Tin’s optimal consumption bundle, \((x^*_t, y^*_t)\).

(e) Is Duncan right that these bundles and these prices make a competitive equilibrium? Explain.

4. Consider Rin and Tin from Question 3. We shall now solve this general equilibrium model. We are free to set one of the prices equal to 1. We will let good \(x\) be the numeraire good; that is, we will set \(p_x = 1\), and we will solve for the appropriate \(p_y\).

(a) Write down Rin’s budget constraint. Solve for Rin’s optimal consumption bundle, \((x^*_r, y^*_r)\), with \(x^*_r\) and \(y^*_r\) figured as functions of \(p_y\).

(b) Write down Tin’s budget constraint. Solve for Tin’s optimal consumption bundle, \((x^*_t, y^*_t)\), \(x^*_t\) and \(y^*_t\) figured as functions of \(p_y\).

(c) Write down the market-clearing (i.e., total demand = total supply) condition for \(x\). Using your answers from (a) and (b), rewrite the market-clearing condition as a function of \(p_y\), and solve for \(p_y\).

(d) Plug \(p_y\) back into your answers from (a) and (b) to find the competitive equilibrium.

5. There are two goods in the world, milk \((x)\) and honey \((y)\), and two consumers, Milne and Shepard. Milne’s utility function is \(u_m = x_m y_m^2\) and his endowment is \(\omega_m = (4, 4)\). Shepard’s utility function is \(u_s = x_s y_s\) and his endowment is \(\omega_s = (0, 0)\). We will again let \(p_x = 1\), and we will let \(p_y\) vary.

(a) Is the original endowment Pareto optimal? Explain. Suppose the dictator sets Milne’s lump-sum tax at \(T_m = -4\), and Shepard’s lump-sum transfer at \(T_s = 4\).

(b) Write down Milne’s new budget constraint. Solve for Milne’s optimal consumption bundle, \((x^*_m, y^*_m)\), with \(x^*_m\) and \(y^*_m\) figured as functions of \(p_y\).

(c) Write down Shepard’s new budget constraint. Solve for Shepard’s optimal consumption bundle, \((x^*_s, y^*_s)\), with \(x^*_s\) and \(y^*_s\) figured as functions of \(p_y\).

(d) Solve for the competitive equilibrium; that is, the market-clearing \((x^*_m, y^*_m)\) and \((x^*_s, y^*_s)\), and the equilibrium \(p_y\).

(e) Prove that the new equilibrium allocation is Pareto optimal.

6. “To achieve an efficient allocation, lump-sum taxes on consumers’ endowments and per unit taxes on the prices of goods are equivalent.” Do you agree with this assertion? Explain, using the welfare theorems in your arguments.