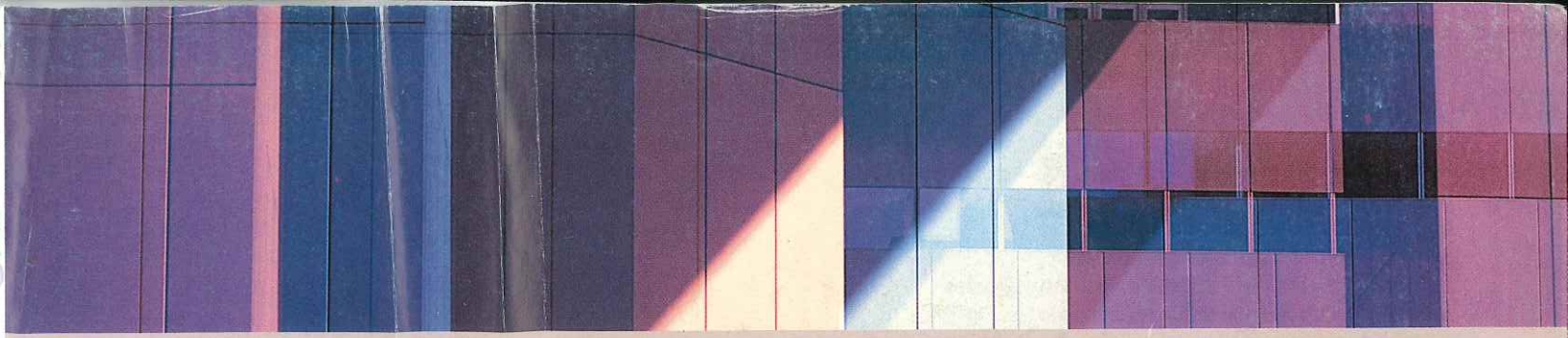
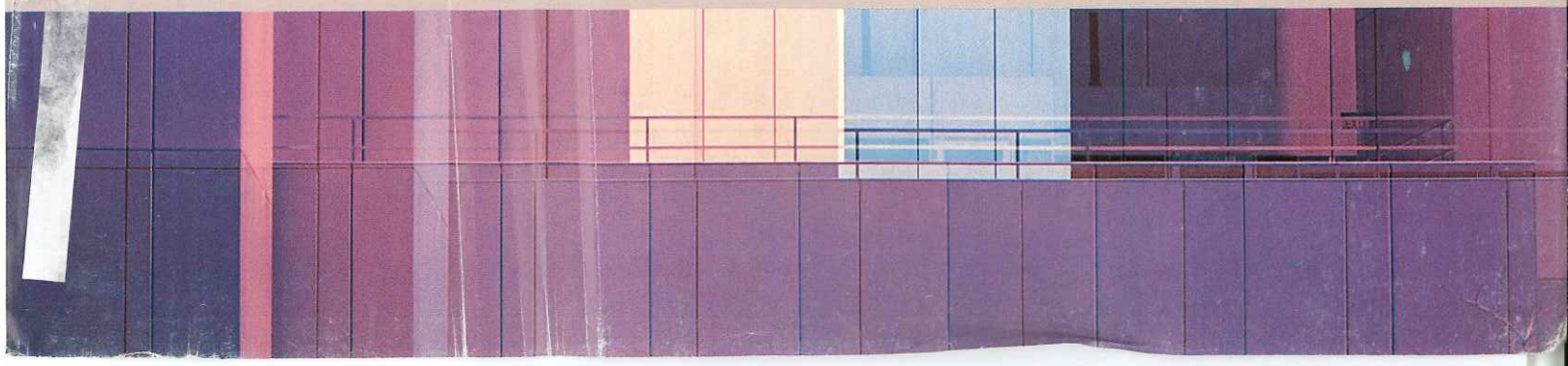


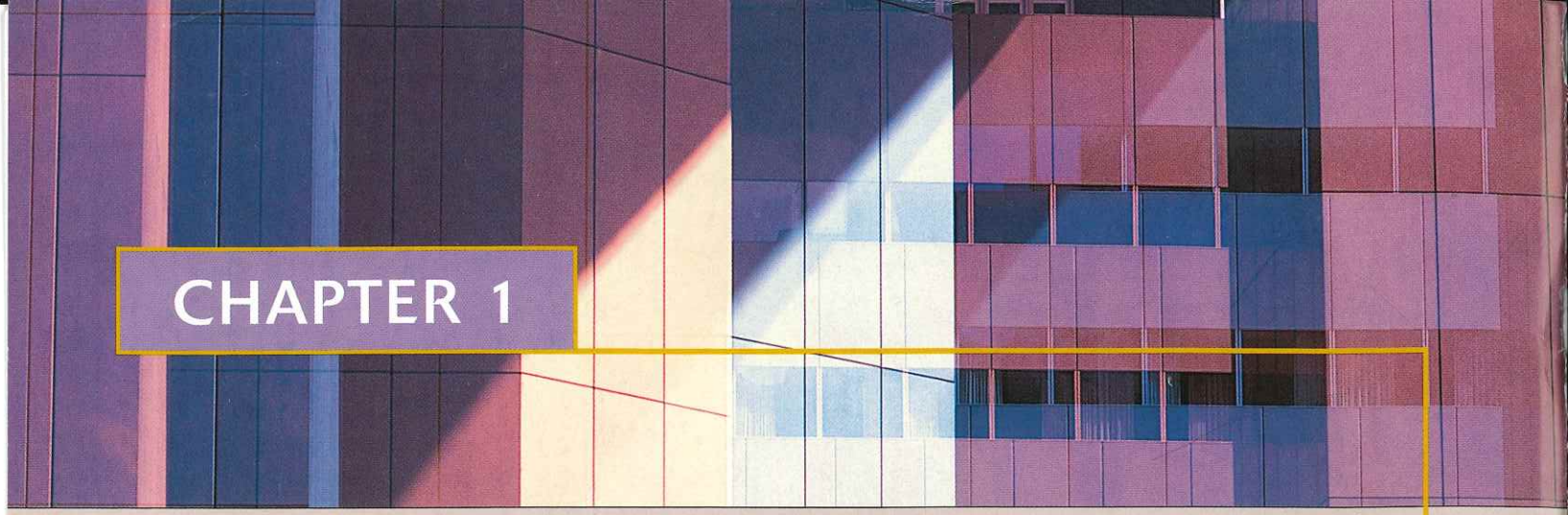
Economics

FIFTH EDITION

John B. Taylor

Houghton Mifflin Company
Boston New York





CHAPTER 1

The Central Idea

This is a true story. In the spring of 1996, a 19-year-old college sophomore, who had just finished taking introductory economics, was faced with a *choice*: to continue college for an additional two years or to devote full time to a job. The job was being a professional golfer on the Pro Tour—a job for which the sophomore was uniquely qualified, having won three U.S. amateur titles. Doing both college and the Pro Tour was not an option because time is *scarce*: With only 24 hours in the day, he simply did not have time to do both. For this sophomore, completing college had a great cost: not only the two years of college expenses, but also the forgone tournament winnings and advertising endorsements that a successful pro golf career would bring. The golfer—his name is Tiger Woods—made a choice. He became a pro. By the fall of 1996 he was selected Sportsman of the Year by *Sports Illustrated*. In 1997 he stunned the golfing world with a record-setting win of the venerable Masters Tournament, and with his second Masters victory in 2001, Woods became the first player to hold all four major professional championships at the same time. He won the Masters again in 2005, by which time he had earned over $50 million in prizes worldwide, and much more in advertising and endorsements.

Tiger Woods would not have had the same opportunity had he not been able to *interact with people*. Golf fans who enjoyed watching him play golf interacted with him: They paid to see him play. Executives who ran companies like Nike and American Express interacted with him: They paid him for his endorsement. And Tiger's family, teachers, and friends interacted with him: They conveyed basic

skills, enhanced his confidence, and helped him remain cool under pressure. Tiger gained from these interactions with people—and the people gained, too.

This true story illustrates the idea that lies at the center of economics: that people make *purposeful choices* with *scarce* resources, and *interact* with other people when they make these choices. More than anything else, **economics** is the study of how people deal with scarcity.

economics: the study of how people deal with scarcity.

scarcity: the situation in which the quantity of resources is insufficient to meet all wants.

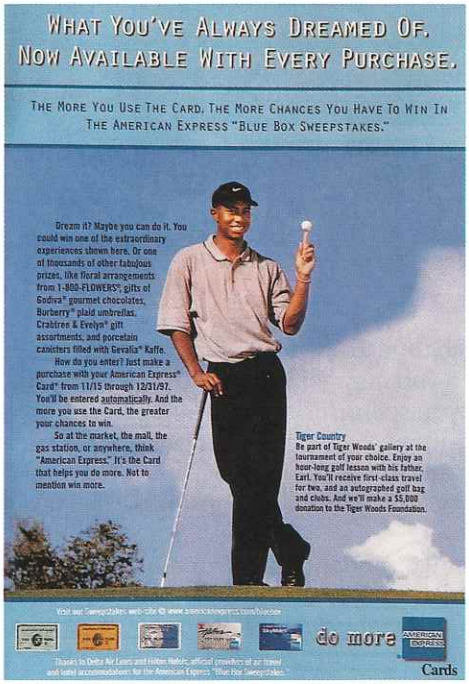
choice: a selection among alternative goods, services, or actions.

economic interaction: exchanges of goods and services between people.

market: an arrangement by which economic exchanges between people take place.

Scarcity is a situation in which people's resources are limited. People always face a scarcity of something—frequently, as in Tiger Woods's case, time. Scarcity implies that people must make a **choice** to forgo, or give up, one thing in favor of another: to work full time or go to school, to take economics or biology this term, to work or stay at home with the children.

Economic interaction between people occurs every time they trade or exchange goods or services with each other. For example, a college student buys education services from a college in exchange for tuition. A teenager sells labor services to Taco Bell in exchange for cash. Within a family, one spouse may agree to cook every day in exchange for the other spouse's doing the dishes every day.



The choice was to continue college or join the Pro Tour. What would you have done?

As these examples indicate, economic interactions can occur either within an *organization* or *group*, such as a family, or in a *market*. A **market** is simply an arrangement by which buyers and sellers can interact and exchange goods and services with each other. Examples of interactions in markets are the buying and selling of a college education and the buying and selling of the labor services of Taco Bell workers. There are many, many other markets, from the New York stock market to a local flea market.

The purpose of this book is to introduce you to the field of economics and to help you understand the economic challenges and opportunities you face as an individual in the economy. The goal is not to peek passively at economics but to learn to think like an economist.

The first step toward this goal is for you to get an intuitive feel for how pervasive scarcity, choice, and economic interaction are in the real world—and to learn some of the powerful implications of this basic fact of economic life. That is the purpose of this chapter.

Scarcity and Choice for Individuals

It is easy to find everyday examples of how people make purposeful choices when they are confronted with a scarcity of time or resources. A choice that may be on your

Gains from Trade on the Internet

The Internet has created many new opportunities for gains from trade. Internet auction sites like eBay allow sellers a way to offer their goods for sale and buyers a way to make bids on sale items. The gains are similar to those of Maria and Adam as they trade sunglasses for hats. Hundreds of different types of sunglasses and baseball hats (and millions of other things) can be bought and sold on eBay—5,566 types of sunglasses and 913 types of baseball hats were for sale at last count.

If you—like Maria—want to sell a pair of sunglasses and buy a hat, you can simply go to *www.ebay.com*, offer a pair of sunglasses to sell, and search for the hat you would like to buy. The computer screen will show photos of some of the sunglasses and baseball hats offered.

Whereas many Internet-related businesses started in the mid-1990s have contracted or disappeared completely in the tech implosion of recent years, eBay has remained hugely successful—perhaps because it is so simple. It provides information and a means of communicating transactions—a sort of virtual flea market. Today there are nearly 50 million registered users, with millions of sales transacted in a single day.

How well do you think the simple example of Maria and Adam illustrates the real-world gains from trade by people using eBay to buy and sell things?

The founder (left) and the chief executive (right) of eBay offer new ways to gain from trade.

mind when you study economics is how much time to spend on it versus other activities. If you spend all your time on economics, you may get 100 on the final exam, but that might mean you get a zero in biology. If you spend all your time on biology, then you may get 100 in biology and a zero in economics. Most people resolve the choice by *balancing* out their time to get a decent grade in both subjects. If you are premed, then biology will probably get more time. If you are interested in business, then more time on economics might be appropriate.

Now let us apply this basic principle to two fundamental economic problems: individual choices about what to *consume* and *produce*. For each type of economic problem, we first show how scarcity forces one to make a choice, then show how people gain from interacting with other people.

Consumer Decisions

Consider Maria, who is going for a hike in a park on a sunny day. Maria would love to wear a hat (baseball style with her school logo) and sunglasses on the hike, but she has brought neither with her. Maria has brought $20 with her, however, and there is a store in the park that is having a "two for one" sale. She can buy two hats for $20 or two pairs of sunglasses for $20. Her scarcity of funds causes her to make a choice. The $20 limit on her spending is an example of a *budget constraint*, because she is limited to spending no more than this amount. Her choice will depend on her tastes. Let us

assume that when she is forced by scarcity to make a choice, she will choose the sunglasses. She would prefer to buy one hat and one pair of sunglasses, but that is not possible.

opportunity cost: the value of the next-best forgone alternative that was not chosen because something else was chosen.

■ **Opportunity Cost.** Maria's decision is an example of an economic problem that all people face: A budget constraint forces them to make a choice between different items that they want. Such choices create opportunity costs. The **opportunity cost** of a choice is the value of the forgone alternative that was not chosen. The opportunity cost of the hats is the loss from not being able to wear the sunglasses. An opportunity cost occurs every time there is a choice. For example, the opportunity cost of going to an 8 A.M. class rather than sleeping in is the sleep you lose when you get up early. The opportunity cost of Tiger Woods's staying in college was millions of dollars in prize money.

In many cases involving choice and scarcity, there are many more than two things to choose from. If you choose vanilla ice cream out of a list of many possible flavors, then the opportunity cost is the loss from not being able to consume the *next-best* flavor, perhaps strawberry. In general, when there are more than two items, the opportunity cost is the value of the next-best alternative.

Now, suppose Maria is not the only hiker. Also in the park is Adam, who also has $20 to spend. Adam also loves both hats and sunglasses, but he likes hats more than sunglasses. When forced to make a choice, he buys the hats. His decision is shaped by scarcity just as Maria's is: Scarcity comes from the budget constraint; he must make a choice, and there is an opportunity cost for each choice.

gains from trade: improvements in income, production, or satisfaction owing to the exchange of goods or services.

■ **Gains from Trade: A Better Allocation.** Now suppose that Adam and Maria meet each other in the park. Let's consider the possibility of economic interaction between them. Maria has two pairs of sunglasses and Adam has two hats, so Maria and Adam can trade with each other. Maria can trade one of her pairs of sunglasses for one of Adam's hats, as shown in Figure 1.1. Through such a trade, both Maria and Adam can improve their situation. There are **gains from trade** because the trade reallocates goods between the two individuals in a way they both prefer. Trade occurs because Maria is willing to exchange one pair of sunglasses for one hat, and Adam is willing to exchange one hat for one pair of sunglasses. Because trade is mutually advantageous for both Maria and Adam, they will voluntarily engage in it if they are able to. In fact, if they do not gain from the trade, then neither will bother to make the trade.

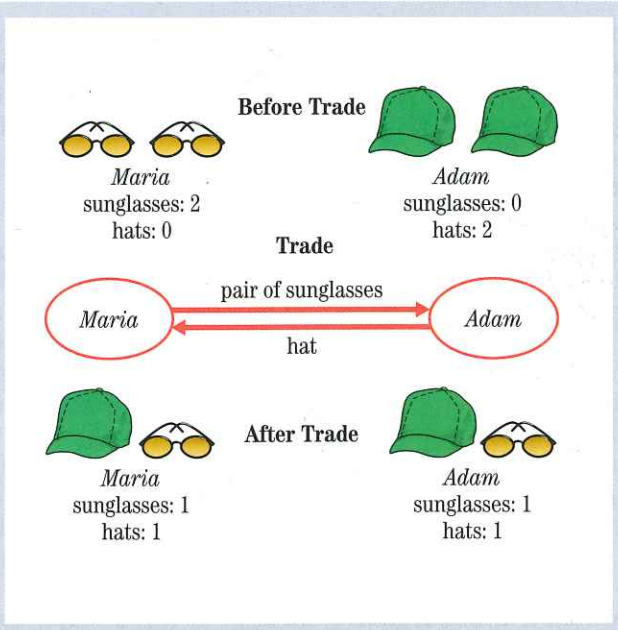


Figure 1.1
Gains from Trade Through a Better Allocation of Goods
Without trade, Maria has more pairs of sunglasses than she would like, and Adam has more hats than he would like. By trading a hat for a pair of sunglasses, they both gain.

This trade is an example of an economic interaction in which a reallocation of goods through trade makes both people better off. There is no change in the total quantity of goods produced. The number of hats and sunglasses has remained the same. Trade simply reallocates existing goods.

The trade between Maria and Adam is typical of many economic interactions we will study in this book. Thinking like an economist in this example means recognizing that a voluntary exchange of goods between people must make them better off. Many economic exchanges are like this, even though they are more complicated than the exchange of hats and sunglasses.

Producer Decisions

Now consider two producers—Emily, a poet, and Johann, a printer. Both face scarcity and must make choices. Because of differences in training, abilities, or inclination, Emily is much better at writing poetry than Johann, but Johann is much better at printing greeting cards than Emily. Suppose that Emily and Johann cannot interact with each other (perhaps they live on different islands and cannot communicate).

If Emily writes poetry full time, she can produce 10 poems in a day; but if she wants to make and sell greeting cards with her poems in them, she must spend some time printing cards and thereby spend less time writing poems. However, Emily is not very good at printing cards; it takes her so much time to do so that if she prints one card, she has time to write only 1 poem rather than 10 poems during the day.

If Johann prints full time, he can produce 10 different greeting cards in a day. However, if he wants to sell greeting cards, he must write poems to put inside them. Johann is so poor at writing poems that if he writes only 1 poem a day, his production of greeting cards drops from 10 to 1 per day.

Following is a summary of the choices Emily and Johann face because of a scarcity of time and resources.

	Emily, the Poet		**Johann, the Printer**	
	Write Full Time	*Write and Print*	*Print Full Time*	*Write and Print*
Cards	0	1	10	1
Poems	10	1	0	1

If Emily and Johann cannot interact, then each can produce only 1 greeting card with a poem on the inside in a day. Alternatively, Emily could produce 10 poems without the cards and Johann could produce 10 cards without the poems, but then neither would earn anything. We therefore assume that when confronted with this choice, both Emily and Johann will each choose to produce 1 greeting card with a poem inside. In total, they produce 2 greeting cards.

■ **Gains from Trade: Greater Production.** Now consider the possibility of economic interaction. Suppose that Emily and Johann can trade. Johann could sell his printing services to Emily, agreeing to print her poems on nice greeting cards. Then Emily could sell the greeting cards to people. Under this arrangement, Emily could spend all day writing poetry, and Johann could spend all day printing. In total, they could produce 10 different greeting cards together, expending the same time and effort it took to produce 2 greeting cards when they could not trade.

Note that in this example the interaction took place in a market: Johann sold his print jobs to Emily. Another approach would be for Emily and Johann to go into business together, forming a firm, Dickinson and Gutenberg Greetings, Inc. Then their economic interaction would occur within the firm, without buying or selling in the market.

Whether in a market or within a firm, the gains from trade in this example are huge. By trading, Emily and Johann can increase their production of greeting cards by five times, from 2 cards to 10 cards.

■ **Specialization, Division of Labor, and Comparative Advantage.** This example illustrates another way in which economic interaction improves people's lives. Economic interaction allows for *specialization*: people concentrating on what

division of labor: the division of production into various parts in which different groups of workers specialize.

they are good at. Emily specializes in poetry, and Johann specializes in printing. The specialization creates a division of labor. A **division of labor** occurs when some workers specialize in one task while others specialize in another task. They divide the overall production into parts, with some workers concentrating on one part (printing) and other workers concentrating on another part (writing).

comparative advantage: a situation in which a person or group can produce one good at a lower opportunity cost than another person or group.

The poetry/printing example of Emily and Johann also illustrates another economic concept, **comparative advantage.** In general, a person or group of people has a comparative advantage in producing one good relative to another good if that person or group can produce that good with comparatively less time, effort, or resources than another person or group can produce that good. For example, compared with Johann, Emily has a comparative advantage in writing relative to printing. And compared with Emily, Johann has a comparative advantage in printing relative to writing. As this example shows, production can be increased if people specialize in the skill in which they have a comparative advantage[1]—that is, if Emily specializes in writing and Johann in printing.

International Trade

Thus far, we have said nothing about where Emily and Johann live or work. They could reside in the same country, but they could also reside in different countries. Emily could live in the United States; Johann, in Germany. If this is so, when Emily purchases Johann's printing service, **international trade** will take place because the trade is between people in two different countries. Similarly, Maria could live in Detroit, Michigan, and Adam in Windsor, Ontario. If this is so, their trade will also be international.

international trade: the exchange of goods and services between people or firms in different nations.

The gains from international trade are thus of the same kind as the gains from trade within a country. By trading, people can better satisfy their preferences for goods (as in the case of Maria and Adam), or they can better utilize their comparative advantage (as in the case of Emily and Johann). In either situation, there is a gain to both participants from trade.

REVIEW

- All individuals face scarcity in one form or another. Scarcity forces people to make choices. When there is choice, there is also an opportunity cost of not doing one thing because another thing has been chosen.
- People benefit from economic interactions—trading goods and services—with other people.
- Gains from trade occur because goods and services can be allocated in ways that are more satisfactory to people.
- Gains from trade also occur because trade permits specialization through the division of labor. People should specialize in the production of goods in which they have a comparative advantage.

1. Other examples are explored in the chapter "The Gains from International Trade," where you can see that comparative advantage can also occur when one person is absolutely better at both activities.

Scarcity and Choice for the Economy as a Whole

Just as individuals face scarcity and choice, so does the economy as a whole. The total amount of resources in an economy—workers, land, machinery, factories—is limited. Thus, the economy cannot produce all the health care, crime prevention, education, or entertainment that people want. A choice must be made. Let us first consider how to represent scarcity and choice in the whole economy and then consider alternative ways to make the choices.

Production Possibilities

Table 1.1
Production Possibilities

	Movies	Computers
A	0	25,000
B	100	24,000
C	200	22,000
D	300	18,000
E	400	13,000
F	500	0

production possibilities: alternative combinations of production of various goods that are possible, given the economy's resources.

To simplify things, let us suppose that production in the economy can be divided into two broad categories. Suppose the economy can produce either computers (mainframes, PCs, hand calculators) or movies (thrillers, love stories, mysteries, musicals). The choice between computers and movies is symbolic of one of the most fundamental choices individuals in any society must face: how much to invest in order to produce more or better goods in the future versus how much to consume in the present. Computers help people produce more or better goods. Movies are a form of consumption. Other pairs of goods could also be used in our example. Another popular example is guns versus butter, representing defense goods versus nondefense goods.

With a scarcity of resources such as labor and capital, there is a choice between producing some goods, such as computers, versus other goods, such as movies. If the economy produces more of one, then it must produce less of the other. Table 1.1 gives an example of the alternative choices, or the **production possibilities,** for computers and movies. Observe that there are six different choices, some with more computers and fewer movies, others with fewer computers and more movies.

Table 1.1 tells us what happens as available resources in the economy are moved from movie production to computer production or vice versa. If resources move from producing movies to producing computers, then fewer movies are produced. For example, if all resources are used to produce computers, then 25,000 computers and zero movies can be produced, according to the table. If all resources are used to produce movies, then no computers can be produced. These are two extremes, of course. If 100 movies are produced, then we produce 24,000 computers rather than 25,000 computers. If 200 movies are produced, then computer production must fall to 22,000.

Increasing Opportunity Costs

The production possibilities in Table 1.1 illustrate the concept of opportunity cost for the economy as a whole. The opportunity cost of producing more movies is the value of the forgone computers. For example, the opportunity cost of producing 200 movies rather than 100 movies is 2,000 computers.

An important economic idea about opportunity costs is demonstrated in Table 1.1. Observe that movie production increases as we move down the table. As we move from row to row, movie production increases by the same number: 100 movies. The decline in computer production between the first and second rows—from 25,000 to 24,000 computers—is 1,000 computers. The decline between the second and third rows—from 24,000 to 22,000 computers—is 2,000 computers. Thus, the decline in computer production gets greater as we produce more movies. As we

Teaching Jobs and Graduate School Applications—Two Sides of the Same Coin

Dozens of new teachers join schools in Silicon Valley, California. Applications to MBA programs at Chicago and MIT soar. Do these two seemingly unrelated events have anything in common? Actually, they do. Behind them we find the same economic phenomenon at work: opportunity costs.

For years, California and other parts of the United States have experienced teacher shortages. During the economic boom of the 1990s, college graduates who might have been interested in teaching had better-paying alternatives in the private sector. In 2000, a teaching job in Silicon Valley paid an average salary of $50,000, while a job in the computer industry paid an average of $80,000, not counting possible gains from stock options—at least a $30,000 differential.

At the same time, college graduates who were considering advancing their education faced a similar decision: "Should I get an MBA and improve my career and future salaries, or should I accept an immediate, high-paying job at a start-up or consulting firm?"

As the recession hit the United States in 2001, many workers were laid off, and others saw their salaries reduced. The U.S. unemployment rate grew from 4.7 percent in January 2001 to 5.6 percent in January 2002, and in Santa Clara County—the heart of Silicon Valley—the increase in the unemployment rate was more dramatic: from 1.7 percent to 7.7 percent. Hewlett-Packard, for example, laid off 6,000 workers—almost 7 percent of its work force—while one of its spinoffs, Agilent, reduced salaries 10 percent for all its 48,000 employees in 2001.

With lower salaries and fewer jobs in the private sector, the opportunity cost of teaching and studying fell. Business schools reported a barrage of applications, with increases of between 50 and 100 percent over the previous year. School districts witnessed a sharp decrease in the number of vacancies available, with many new teachers willing to undergo months of training and substantial pay cuts relative to their old high-tech jobs.

Think of the options you will be facing when you graduate. Given the jobs and salaries currently available, what career do you think you would like to pursue? What would your opportunity cost be?

move from 400 movies to 500 movies, we lose 13,000 computers. In other words, the opportunity cost, in terms of computers, of producing more movies increases as we produce more movies. Each extra movie requires a loss of more and more computers. What we have just described is called **increasing opportunity costs,** with emphasis on the word *increasing*.

increasing opportunity cost: a situation in which producing more of one good requires giving up an increasing amount of production of another good.

Why do opportunity costs increase? You can think about it in the following way. Some of the available resources are better suited for movie production than for computer production, and vice versa. Workers who are good at building computers might not be so good at acting, for example, or moviemaking may require an area with a dry, sunny climate. As more and more resources go into making movies, we are forced to take resources that are much better at computer making and use them for moviemaking. Thus, more and more computer production must be lost to increase the production of movies by the same amount. Adding specialized computer designers to a movie cast would be very costly in terms of lost computers, and it might add little to movie production.

The Production Possibilities Curve

Figure 1.2 is a graphical representation of the production possibilities in Table 1.1 that nicely illustrates increasing opportunity costs. We put movies on the horizontal axis and computers on the vertical axis of the figure. Each pair of numbers in a row of the table becomes a point on the graph. For example, point *A* on the graph is from row A of the table. Point *B* is from row B, and so on.

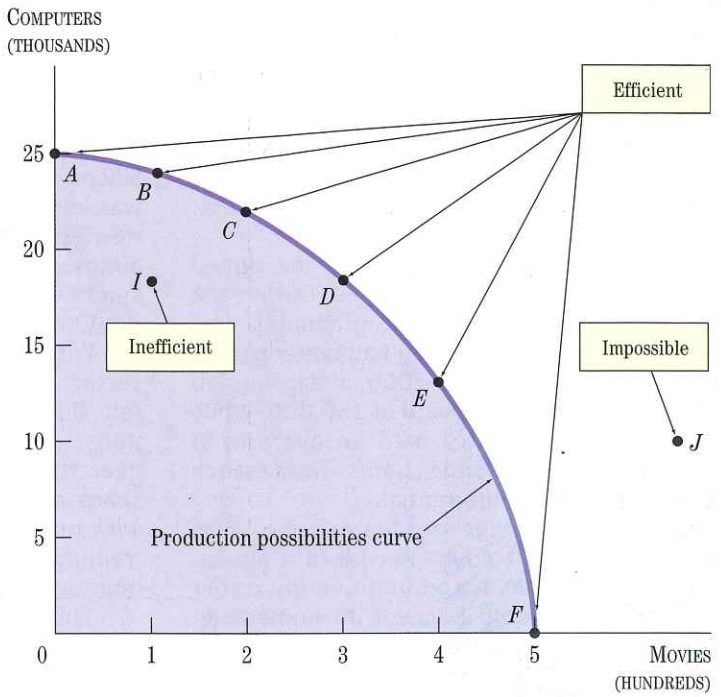


Figure 1.2
The Production Possibilities Curve
Each point on the curve shows the maximum amount of computers that can be produced when a given amount of movies is produced. The points with letters are the same as those in Table 1.1 and are connected by smooth lines. Points in the shaded area inside the curve are inefficient. Points outside the curve are impossible. For the efficient points on the curve, the more movies that are produced, the fewer computers that are produced. The curve is bowed out because of increasing opportunity costs.

production possibilities curve: a curve showing the maximum combinations of production of two goods that are possible, given the economy's resources.

When we connect the points in Figure 1.2, we obtain the **production possibilities curve.** It shows the maximum number of computers that can be produced for each quantity of movies produced. Note that the curve in Figure 1.2 slopes downward and is bowed out from the origin. That the curve is bowed out indicates that the opportunity cost of producing movies increases as more movies are produced. As resources move from computer making to moviemaking, each additional movie means a greater loss of computer production.

■ **Inefficient, Efficient, or Impossible?** The production possibilities curve shows the effects of scarcity and choice in the economy as a whole. Three situations can be distinguished in Figure 1.2, depending on whether production is in the shaded area, on the curve, or outside the curve.

First, imagine production at point *I*. This point, with 100 movies and 18,000 computers, is inside the curve. But the production possibilities curve tells us that it is possible with the same amount of resources to produce more computers, more movies, or both. For some reason, the economy is not working well at point *I*. For example, instead of using movie film, people may be taking still photos and then sticking them together with tape to make the movie. Points inside the curve, like point *I*, are *inefficient* because the economy could produce a larger number of movies, as at point *D*, or a larger number of computers, as at point *B*. Points inside the production possibilities curve are possible, but they are inefficient.

Second, consider points on the production possibilities curve. These points are *efficient*. They represent the maximum amount that can be produced with available resources. The only way to raise production of one good is to lower production of the other good. Thus, points on the curve show a *tradeoff* between one good and another.

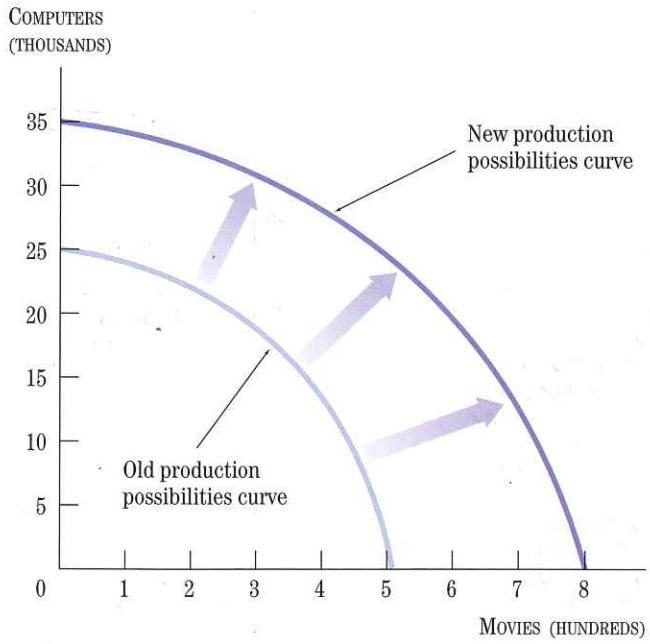


Figure 1.3
Shifts in the Production Possibilities Curve
The production possibilities curve shifts out as the economy grows. The maximum amounts of movies and computers that can be produced increase. Improvements in technology, more machines, or more labor permit the economy to produce more.

Third, consider points to the right and above the production possibilities curve, like point *J* in Figure 1.2. These points are *impossible*. The economy does not have the resources to produce those quantities.

■ **Shifts in the Production Possibilities Curve.** The production possibilities curve is not immovable. It can *shift* out or in. For example, the curve is shown to shift out in Figure 1.3. More resources—more workers, for example—shift the production possibilities curve out. A technological innovation that allowed one to edit movies faster would also shift the curve outward. So would more cameras, lights, and studios. When the production possibilities curve shifts out, the economy grows because more goods and services can be produced. The production possibilities curve need not shift outward by the same amount in all directions. There can be more movement up than to the right, for example.

As the production possibilities curve shifts out, impossibilities are converted into possibilities. Some of what was impossible for the U.S. economy in 1970 is possible now. Some of what is impossible now will be possible in 2020. Hence, the economists' notion of possibilities is a temporary one. When we say that a certain combination of computers and movies is impossible, we do not mean "forever impossible," we only mean "currently impossible."

■ **Scarcity, Choice, and Economic Progress.** However, the conversion of impossibilities into possibilities is also an economic problem of choice and scarcity: If we invest less now—in machines, in education, in children, in technology—and consume more now, then we will have less available in the future. If we take computers and movies as symbolic of investment and consumption, then choosing more

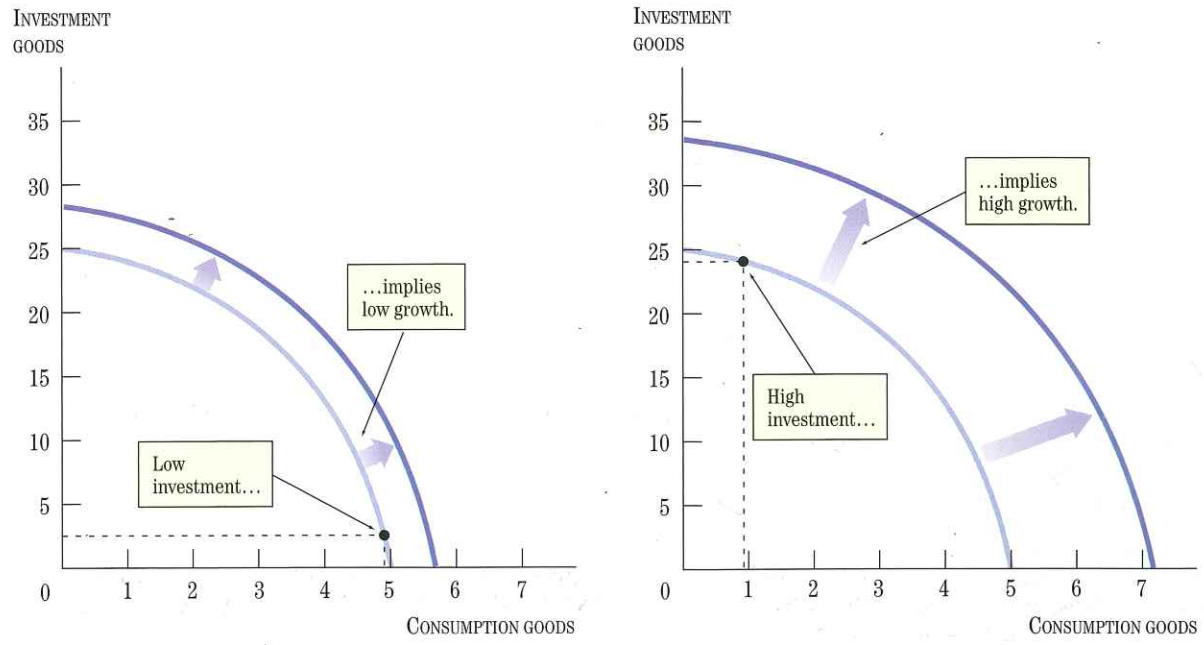


Figure 1.4
Shifts in the Production Possibilities Curve Depend on Choices

On the left, few resources are devoted to investment for the future; hence, the production possibilities curve shifts only a little over time. On the right, more resources are devoted to investment and less to consumption; hence, the production possibilities curve shifts out by a larger amount over time.

investment will result in a larger outward shift of the production possibilities curve, as illustrated in Figure 1.4. More investment enables the economy to produce more in the future.

The production possibilities curve represents a *tradeoff*, but it does not mean that some people win only if others lose. First, it is not necessary for someone to lose in order for the production possibilities curve to shift out. When the curve shifts out, the production of both items increases. Although some people may fare better than others as the production possibilities curve is pushed out, no one necessarily loses. In principle, everyone can gain. Second, if the economy is at an inefficient point (like point *I* in Figure 1.2), then production of both goods can be increased with no tradeoff. In general, therefore, the economy is more like a win-win situation, where everyone can achieve a gain.

REVIEW

- The production possibilities curve represents the choices open to a whole economy when it is confronted by a scarcity of resources. As more of one item is produced, less of another item must be produced. The opportunity cost of producing more of one item is the reduced production of another item.
- The production possibilities curve is bowed out because of increasing opportunity costs.

- Points inside the curve are inefficient. Points on the curve are efficient. Points outside the curve are impossible.
- The production possibilities curve shifts out as resources increase.
- Outward shifts of the production possibilities curve or moves from inefficient to efficient points are the reasons why the economy is not a zero-sum game, despite the existence of scarcity and choice.

Market Economies and the Price System

The Three Fundamental Economic Questions
Any economic system has to answer three questions: What goods and services should be produced—cars, televisions, or something else? How should these goods or services be produced—in what type of factory, and with how much equipment and labor? And for whom should these goods be produced?

The production possibilities curve enables us to discuss key questions about the economy.

Three Questions

There are three essential questions or problems that every economy must find a way to solve, whether it is a small island economy or a large economy like the United States.

- *What* is to be produced: movies, computers, guns, butter, greeting cards, Rollerblades, health care, or something else? In other words, where on the production possibilities curve should an economy be?

What?

How?

For Whom?

- *How* are these goods to be produced? In other words, how can an economy use the available resources so that it is not at an inefficient point inside the production possibilities curve?
- *For whom* are the goods to be produced? We know from the hat/sunglasses example that the allocation of goods in an economy affects people's well-being. An economy in which Maria could not trade her sunglasses for a hat would not work as well as one in which such trades and reallocations are possible. Moreover, an economy in which some people get everything and others get virtually nothing is also not working well.

market economy: an economy characterized by freely determined prices and the free exchange of goods and services in markets.

command economy: an economy in which the government determines prices and production; also called a centrally planned economy.

Broadly speaking, the **market economy** and the **command economy** are two alternative approaches to answering these questions. In a market economy, most decisions about what, how, and for whom to produce are made by individual consumers, firms, governments, and other organizations interacting in markets. In a command, or centrally planned, economy, most decisions about what, how, and for whom to produce are made by those who control the government, which, through a central plan, commands and controls what people do.

Command economies are much less common today than they were in the mid-twentieth century, when nearly half the world's population, including the residents of Eastern Europe, the Soviet Union, and China, lived in centrally planned economies. After many decades of struggling to make this system work, leaders of the command economies gradually grew disillusioned with the high degree of inefficiency resulting from the planned approach—which required that the state, or central planners, make critical detailed production decisions—which often resulted in shortages or surplus of products and, as a by-product, in political unrest. In recent years, most command economies have been engaged in the complex process of converting from a command to a market system, with varying degrees of success. This is partly due to the fact that these economies had none or few of the social, legal, or political fixtures critical to the market system. China has probably been the most successful of these economies at making the transition, developing a model that the Chinese term *market socialism*. Beginning in the 1970s, elements of both the command and market economies coexisted in China; in the mid 1990s, market mechanisms grew more dominant. While its political system is still highly centralized, many people credit China's rapid economic growth in recent years to its decentralized economic system.

Key Elements of a Market Economy

Let's take a closer look at some of the ingredients critical to a market economy.

freely determined price: a price that is determined by the individuals and firms interacting in markets.

■ **Freely Determined Prices.** In a market economy, most prices—such as the price of computers—are freely determined by individuals and firms interacting in markets. These **freely determined prices** are an essential characteristic of a market economy. In a command economy, most prices are set by government, and this leads to inefficiencies in the economy. For example, in the Soviet Union, the price of bread was set so low that farmers fed bread to the cows. Feeding bread to livestock is an enormous waste of resources. Livestock could eat plain grain. By feeding the cows bread, farmers added the cost of the labor to bake the bread and the fuel to heat the bread ovens to the cost of livestock feed. This is inefficient, like point *I* in Figure 1.2.

In practice, not all prices in market economies are freely determined. For example, some cities control the price of rental apartments. We will look at these exceptions later. But the vast majority of prices are free to vary.

property rights: rights over the use, sale, and proceeds from a good or resource.

■ **Property Rights and Incentives.** Property rights are another key element of a market economy. **Property rights** give individuals the legal authority to keep or sell property, whether land or other resources. Property rights are needed for a market economy because they give people the ability to buy and sell goods. Without property rights, people could take whatever they wanted without paying. People would have to devote time and resources to protecting their earnings or goods.

incentive: a device that motivates people to take action, usually so as to increase economic efficiency.

Moreover, by giving people the rights to the earnings from their work, as well as letting them suffer some of the consequences or losses from their mistakes, property rights provide **incentives.** For example, if an inventor could not get the property rights to an invention, then the incentive to produce the invention would be low or even nonexistent. Hence there would be few inventions, and we would all be worse off. If there were no property rights, people would not have incentives to specialize and reap the gains from the division of labor. Any extra earnings from specialization could be taken away.

■ **Freedom to Trade at Home and Abroad.** Economic interaction is a way to improve economic outcomes, as the examples in this chapter indicate. Allowing people to interact freely is thus another necessary ingredient of a market economy. Freedom to trade can be extended beyond national borders to other economies.

International trade increases the opportunities to gain from trade. This is especially important in small countries, where it is impossible to produce everything. But the gains from exchange and comparative advantage also exist for larger countries.

■ **A Role for Government.** Just because prices are freely determined and people are free to trade in a market economy does not mean that there is no role for government. For example, in virtually all market economies, the government provides defense and police protection. The government also helps establish property rights. But how far beyond that should it go? Should the government also address the "for whom" question by providing a safety net—a mechanism to deal with the individuals in the economy who are poor, who go bankrupt, who remain unemployed? Most would say yes, but what should the government's role be? Economics provides an analytical framework to answer such questions. In certain circumstances—called **market failure**—the market economy does not provide good enough answers to the "what, how, and for whom" questions, and the government has a role to play in improving on the market. However, the government, even in the case of market failure, may do worse than the market, in which case economists say there is **government failure.**

market failure: any situation in which the market does not lead to an efficient economic outcome and in which there is a potential role for government.

government failure: a situation in which the government makes things worse than the market, even though there may be market failure.

■ **The Role of Private Organizations.** It is an interesting feature of market economies that many economic interactions between people take place in organizations—firms, families, charitable organizations—rather than in markets. Some economic interactions that take place in organizations could take place in the market. In some circumstances, the same type of interaction takes place in a firm and in a market simultaneously. For example, many large firms employ lawyers as part of their permanent staff. Other firms simply purchase the services of such lawyers in the market; if the firm wants to sue someone or is being sued by someone, it hires an outside lawyer to represent it.

Economic interactions in firms differ from those in the market. Staff lawyers inside large firms are usually paid annual salaries that do not depend directly on the number of hours worked or their success in the lawsuits. In contrast, outside lawyers are paid an hourly fee and a contingency fee based on the number of hours worked and how successful they are.

Economic interactions can take place in a marketplace, between two friends, on the Internet, or even on the radio, as this article attests. All you need is someone who wants to sell something and someone else who wants to buy. Tradio helps potential buyers and sellers learn about each other in a virtual market, realizing gains from trade that would not be possible without the exchange of information. Even though radio stations do not charge buyers and sellers, you will notice that tradio is not just a public service. Tradio is very low-cost programming that attracts a large audience and many paid ads from local businesses—a win-win situation for individual buyers and sellers, radio stations, and local businesses. Gains from trade in action!

Who Needs eBay? For Towns Across U.S., Tradio Is Real Deal

By Reid J. Epstein, Staff Reporter
THE WALL STREET JOURNAL
September 11, 2002

GLASGOW, Mont.—It was a little after nine one recent morning, and local residents were already on the line to Lori Mason's radio show.

One caller wanted to unload a riding lawnmower ($500). Another tried to sell an irrigation pump ($100), and four offered up washing machines, including one that "leaks a little bit" ($10). Others still were looking to buy eight bales of straw, fresh dill and a large dog house.

When one young woman phoned in to put her '79 GMC pickup on the block, the 54-year-old Ms. Mason not only recognized her voice but urged her to loosen up.

"Oh, sorry," said the caller, 23-year-old day-care provider Jamie Seyfert. "We're willing to trade for guns, jet skis or anything fun."

This is the sound of "tradio" (pronounced TRADE-ee-o), a kind of on-air swap meet that has been a fixture of small-town stations from Florida to Alaska for decades. Far from being rendered obsolete by the Internet, many tradio shows are doing surprisingly well these days. They are the top moneymakers for some stations, often commanding a premium from local advertisers. And the format may be pushing into bigger markets. In April, WCCO in Minneapolis introduced "The WCCO Great Garage Sale" and saw its ratings jump 29% in the time slot.

Three-Stoplight Town

Here in Glasgow, a three-stoplight railroad town of 3,253 people on the lonely plains of northeastern Montana, the half-hour show is simulcast three mornings a week from the second-floor studio of locally owned stations KLTZ and KLAN. Virtually everyone in town, from the mayor to the editor of the weekly paper, has bought or sold something on the show.

Here's how it works: Callers announce they're selling something—a gas heater that "would be good in your garage, your huntin' shack or whatever," a "very large collection of Fiestaware dishes in all the new colors" or some "very friendly young goats"—and leave their phone number. Anyone interested calls the seller, and the transaction is negotiated face-to-face.

Internet Connection

While it may sound archaic in the age of eBay, the tradio format seems to be benefiting from the buzz generated by the popular Web auction site. Some tradio shows are using the Internet to their advantage—allowing listeners to submit items for sale via e-mail and posting items called into the show on their Web sites.

The tradio format first took hold in the early 1950s when powerful nationwide radio networks cut back on programming. To fill the void, small-town stations began broadcasting obituaries, birthdays and anniversaries. An appliance-store owner in Seguin is believed by many in the industry to have started the first tradio show. He began buying air time to read notices of items for sale that customers had posted on a bulletin board inside his store.

In small towns that don't have daily newspapers—the closest daily to Glasgow is the *Herald* in Williston, N.D., 144 miles away—tradio takes the place of classifieds and, perhaps more important, gossip. When Ms. Mason heard a caller say he was selling his $100 irrigation pump, she exclaimed, "You're on the new water line!" and quizzed him about the difference that had made in his water supply.

Another big attraction is the price. A classified ad in the weekly *Glasgow Courier* costs $5.25 per column inch, while calls to "Tradio" are free. And eBay, which exacts a sliding fee based on the price of the item sold, also requires a hookup to the Internet.

Alicia Sibley, a 24-year-old hay farmer, is a regular listener to the Glasgow show. She was driving her tractor a few weeks ago when she heard a caller offering to sell a 6-foot freezer. She phoned the seller—who had bought the freezer to stock up on frozen foods for fear of a catastrophe at the turn of the millennium—and made a deal for $200. "It's a real nice one, too," Ms. Sibley said. "I saved around $200."

Incentives within an organization are as important as incentives in markets. If the lawyers on a firm's legal staff get to keep some of the damages the firm wins in a lawsuit, they will have more incentive to do a good job. Some firms even try to create marketlike competition between departments or workers in order to give more incentives.

Why do some economic interactions occur in markets and others in organizations? Ronald Coase of the University of Chicago won the Nobel Prize for showing that organizations such as firms are created to reduce market *transaction costs*, the costs of buying and selling, which include finding a buyer or a seller and reaching agreement on a price. When market transaction costs are high, we see more transactions taking place within organizations. For example, a firm might have a legal staff rather than outside lawyers because searching for a good lawyer every time there is a lawsuit is too costly. In a crisis, a good lawyer may not be available.

The Price System

The previous discussion indicates that in market economies, freely determined prices are essential for determining what is produced, how, and for whom. For this reason, a market economy is said to use *the price system* to solve these problems. In this section, we show that prices do a surprising amount of work: (1) Prices serve as *signals* about what should be produced and consumed when there are changes in tastes or changes in technology, (2) prices provide *incentives* to people to alter their production or consumption, and (3) prices affect the *distribution of income*, or who gets what in the economy.

Let's use an example. Suppose that there is a sudden new trend for college students to ride bicycles more and drive cars less. How do prices help people in the economy decide what to do in response to this new trend?

■ **Signals.** First, consider how the information about the change in tastes is signaled to the producers of bicycles and cars. As students buy more bicycles, the price of bicycles rises. A higher price will signal that it is more profitable for firms to produce more bicycles. In addition, some bicycle components, like lightweight metal, will also increase in price. Increased lightweight metal prices signal that production of lightweight metal should increase. As the price of metal rises, wages for metalworkers may increase. Thus, prices are a signal all the way from the consumer to the metalworkers that more bicycles should be produced. This is what is meant by the expression "prices are a signal."

It is important to note that no single individual knows the information that is transmitted by prices. Any economy is characterized by limited information, where people cannot know the exact reasons why prices for certain goods rise or fall. Hence, it is rather amazing that prices can signal this information.

■ **Incentives.** Now let's use this example to consider how prices provide incentives. A higher price for bicycles will increase the incentives for firms to produce bicycles. Because they receive more for each bicycle, they produce more. If there is a large price increase that is not merely temporary, new firms may enter the bicycle business. In contrast, the reduced prices for cars signal to car producers that production should decrease.

■ **Distribution.** How do prices affect the distribution of income? On the one hand, workers who find the production of the good they make increasing because of the higher demand for bicycles will earn more. On the other hand, income will be

reduced for those who make cars or who have to pay more for bicycles. Local delivery services that use bicycles will see their costs increase.

REVIEW

- The market economy and the command economy are two alternative systems for addressing the questions any economy must face: what to produce, how to produce, and for whom to produce.
- A market economy is characterized by several key elements, such as freely determined prices, property rights, and freedom to trade at home and abroad.
- For a market economy to work well, markets should be competitive and the government should play a role.
- Prices are signals, they provide incentives, and they affect the distribution of income.

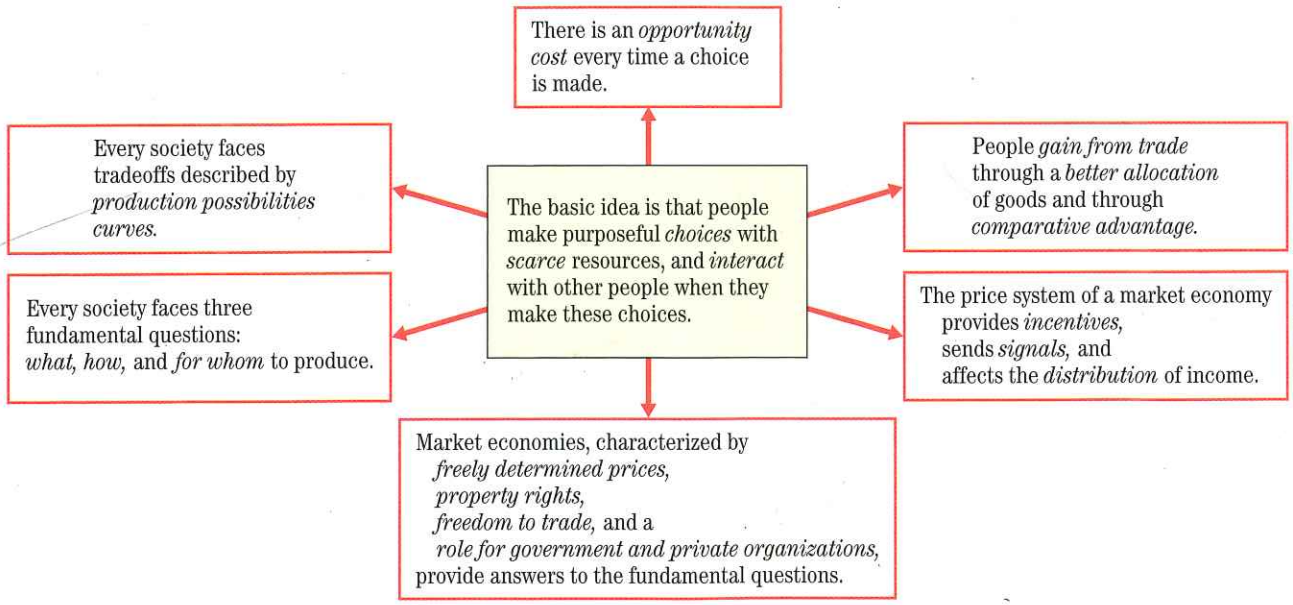


Figure 1.5
From One Central Idea, Many Powerful Ideas Follow

As you study economics, you will see the same central idea again and again. This figure illustrates how many powerful economic ideas are connected to the one in the center.

Conclusion

One basic idea lies at the center of economics: People make purposeful choices with scarce resources, and interact with other people when they make these choices.

This introductory chapter illustrates this idea, starting with Tiger Woods's decision to leave school and continuing with simple examples of people making choices about what to consume or produce.

From this central idea, many other powerful ideas follow, as summarized visually in Figure 1.5. There is an *opportunity cost* every time a choice is made. People *gain from trade*, both through a *better allocation* of goods and through *comparative advantage*. Every society faces tradeoffs described by *production possibilities curves*. Every society faces three fundamental questions: *what*, *how*, and *for whom* to produce. Market economies—characterized by *freely determined prices*, *property rights*, *freedom to trade*, and *a role for both government and private organizations*—give an answer to these three questions. The price system of a market economy works by providing *incentives*, sending *signals*, and affecting the *distribution* of income.

You will see this central idea again and again as you study economics.

KEY POINTS

1. Everyone faces a scarcity of something, usually time or resources.
2. Scarcity leads to choice, and choice leads to opportunity costs.
3. Trade leads to gains because it allows goods and services to be reallocated in a way that improves people's well-being.
4. Trade also leads to gains because it permits people to specialize in what they are relatively good at.
5. The production possibilities curve summarizes the tradeoffs in the whole economy due to scarcity. Economic production is efficient if the economy is on the production possibilities curve.
6. The three basic questions that any economy must face are what, how, and for whom production should take place.
7. A well-functioning market system involves freely determined prices, property rights, freedom to trade, and a role for government and private organizations.
8. If prices are set at the wrong levels by government, waste and inefficiency—such as feeding bread to livestock—will result.
9. Prices transmit signals, provide incentives, and affect the distribution of income.

KEY TERMS

economics
scarcity
choice
economic interaction
market
opportunity cost
gains from trade
division of labor
comparative advantage
international trade
production possibilities
increasing opportunity cost
production possibilities curve
market economy
command economy
freely determined prices
property rights
incentives
market failure
government failure

QUESTIONS FOR REVIEW

1. How do scarcity, choice, and economic interaction fit into the basic idea at the center of economics?
2. Why does scarcity imply a choice among alternatives?
3. Why does choice create opportunity costs?
4. What is the difference between economic interaction in markets and in organizations?
5. Why is there a gain from trade even if total production of goods and services does not change?

6. How can specialization lead to a gain from trade?
7. What is the principle of increasing opportunity costs?
8. What are the key ingredients of a market economy?
9. What are the three basic questions that any economic system must address?
10. What are the three roles of prices?

PROBLEMS

1. Suppose that you are president of the student government, and you have $10,000 for guest speakers for the year. Spike Lee costs $10,000 per appearance. Former economic advisers to the government charge $1,000 per lecture. Hence, you cannot have both Spike Lee and the former economic advisers. Explain the economic problem of choice and scarcity in this case. What issues would you consider in arriving at a decision?
2. Compare the opportunity cost of one more year of school versus working for one year for (1) a high school graduate, (2) a college graduate, and (3) a medical school graduate.
3. Allison will graduate from high school next June. She has ranked her three possible postgraduation plans in the following order: (1) work for two years at a consulting job in her home town paying $20,000 per year, (2) attend a local community college for two years, spending $5,000 per year on tuition and expenses, and (3) travel around the world tutoring a rock star's child for pay of $5,000 per year. What is the opportunity cost of her choice?
4. Suppose you have two salt shakers and a friend of yours has two pepper shakers. Explain how you can both gain from trade. Is this a gain from trade through *better allocation* or *greater production*? Suppose now that your friend lives in another country, whose government does not allow trade between your countries. Who would lose as a result of this trade barrier?
5. Suppose Tina and Julia can produce jars of salsa and computer-designed advertisements in the following combinations in a given week:

Tina		Julia	
Salsa	*Ads*	*Salsa*	*Ads*
50	0	25	0
40	1	20	1
30	2	15	2
20	3	10	3
10	4	5	4
0	5	0	5

   a. If Tina and Julia are each currently producing 2 advertisements per week, how many jars of salsa are they producing? What is the total production of salsa and advertisements between them?
   b. Is there a possibility for increasing production? Why or why not?
   c. Suppose Julia completely specializes in producing advertisements and Tina completely specializes in producing salsa. What will be the total production of advertisements and salsa?
6. Suppose you must divide your time between studying for your math final and writing a final paper for your English class. The fraction of time you spend studying math and its relation to your grade in the two classes is given in the table below.

Fraction of Time Spent on Math	Math Grade	English Grade
0	0	97
20	45	92
40	65	85
60	75	70
80	82	50
100	88	0

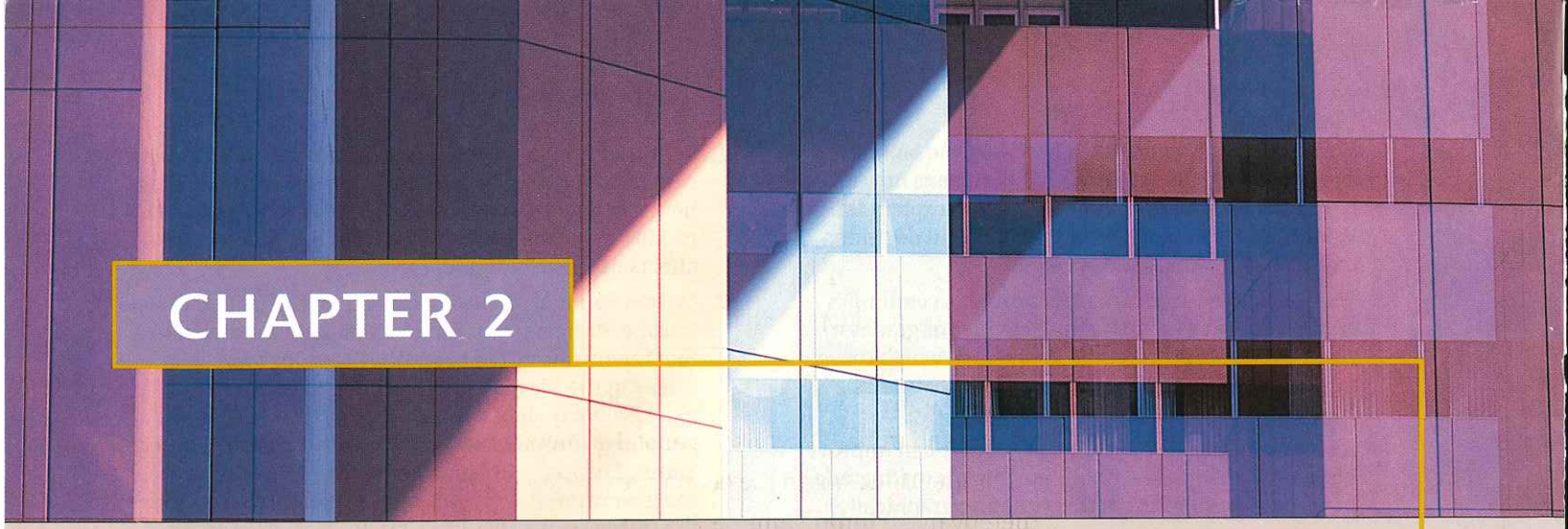
   a. Draw a tradeoff curve for the math grade versus the English grade.
   b. What is the opportunity cost of increasing the time spent on math from 80 to 100 percent? What is the opportunity cost of increasing the time spent on math from 60 to 80 percent?
   c. Are there increasing opportunity costs from spending more time on math? Explain.
   d. What can you do to get a 92 in both subjects? Explain.
7. A small country produces only two goods, cars and computers. Given its limited resources, this country has the following production possibilities:

Cars	Computers
0	200
25	180
50	130
75	70
100	0

   a. Draw the production possibilities curve.
   b. Suppose car production uses mainly machines and computer production uses mainly labor. Show what happens to the curve when the number of machines increases but labor remains unchanged.

8. After World War II and until the 1980s, the Japanese economy grew very rapidly, and its citizens now enjoy a high standard of living. How would you explain the fast Japanese growth in terms of tradeoffs and decisions made by individuals and societies?
9. Tracy tells Huey that he can improve his economics grades without sacrificing fun activities or grades in other courses. Can you imagine ways in which this might be possible? What does that imply about the initial situation?

   If Huey is taking just two courses and he can improve his economics grade without hurting his math grade, how could you represent this situation graphically?
10. Compare two countries. In one country, the government sets prices and never adjusts them. In the other country, the government adjusts prices daily, endeavoring to allocate resources to consumers to satisfy their tastes. Is either of these a market economy? Why or why not?
11. Suppose decreased production of oil in the Middle East causes the price of oil to rise all over the world. Explain how this change in the price signals information to U.S. producers, provides incentives to U.S. producers, and affects the distribution of income.
12. "When you look at the economies in the United States, Europe, or Japan, you see most of the ingredients of a market economy. For example, consider bicycles. Prices in the bicycle market are free to vary; people have property rights to the bicycles they buy; many people sell bicycles; many bicycles sold in the United States, Europe, and Japan come from other countries; the government regulates bicycle use (no bicycles on the freeways, for example); and bicycle production takes place within firms with many workers." Replace bicycles with another good or service of your choosing in this quotation and comment on whether the quotation is still true.



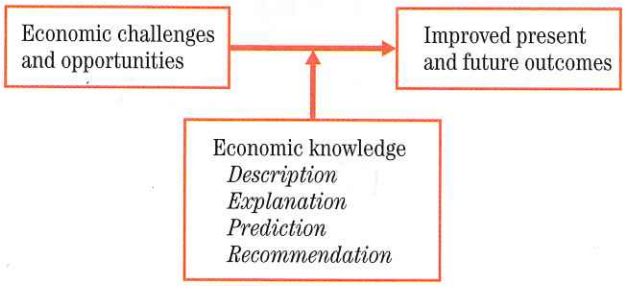
CHAPTER 2

Observing and Explaining the Economy

Mark McClellan is a lot like Tiger Woods. He trained long and hard before he became a professional, and he now excels in his field. But Mark is not an expert at driving and putting golf balls. Rather, he is an expert at observing and explaining economic trends in the biggest industry in America.

After college, Mark went on to earn a Ph.D. in economics. He also went to medical school, completed an internship and residency in a hospital, and became a practicing medical doctor. Why would an economist become a doctor? Because the biggest industry in America is not automobiles, oil, or construction—but health care. Mark concluded that to observe and explain the health-care industry, he needed to be an expert in economics as well as medicine. And all of the training has paid off. As a health-care economist, Mark has demonstrated, among other things, that doctors sometimes provide treatments of little actual value to their patients and that it is the fear of malpractice suits that motivates doctors to do so. His economic expertise and insights on health-care policy have earned him a place, first, in 2001 and 2002, as one of the three members of President Bush's Council of Economic Advisers, then, from 2002 to 2004, as Commissioner of the Food and Drug Administration, and most recently as Administrator of the Centers for Medicare and Medicaid Services.

The purpose of this chapter is to give you a broad overview of economics by looking at the kinds of things economists such as Mark McClellan actually do. Economics is a way of thinking. It entails accurately *describing* economic events, *explaining* why the events occur, *predicting* under what circumstances such events might take place in the future, and *recommending* appropriate courses of action. To make use of economics, you will want to learn to do the describing, the explaining, and even the predicting and recommending yourself—that is, to reason and think like an economist. By making use of economics in this way, you can better understand the economic challenges and opportunities you face, and thereby make improvements in your own life or the lives of those around you.



Just as physicists try to explain the existence of black holes in outer space or biologists try to explain why dinosaurs became extinct, economists try to explain puzzling observations and facts about the economy. Many observations come from everyday life. Are there some economic observations—from your own experience, from recent news stories, or from your history or political science courses—that you find puzzling? Some of your questions might be like these:

- Why is college tuition so high?
- Why have the wages of college graduates increased much more rapidly than the wages of high school dropouts since the 1970s?
- Why are there 17 different types of Colgate toothpaste now, while there were only two types 30 years ago?
- Why is the average income of people in the United States about 35 times higher than that of people in China?
- Why is unemployment much higher in Europe than in the United States?
- Why has health-care spending increased faster than the rest of the U.S. economy?
- Why has the price of health care increased more than most other prices?

All these questions are based on observations about the economy. Some, like the question about college tuition, are fairly obvious and are based on casual observation. But in order to answer such questions, economists, like physicists

or biologists, need to systematically document and quantify their observations and look for patterns. If we can establish the date when dinosaurs became extinct, then we may be able to test our hunch that a cataclysmic event such as an asteroid hitting the earth caused their extinction. To illustrate how economists document and quantify their observations, let us briefly focus on the last two of the preceding questions, the ones concerning health care and the economy. This will also give us an opportunity to introduce some key indicators used to measure the economy.

Health-Care Spending in America

Health-care spending and prices are more than a curiosity. The more a society spends on health care, the less it can spend on other things, as we know from the production possibilities curve of Chapter 1. Concerns about health-care spending have led to major proposals for changing the way health care is provided and paid for in the United States. Health care *is* a major political issue. People who are dissatisfied with their own health care want the right to sue their health-care providers. Debates about the addition of a prescription drug benefit to Medicare—a government health-care program for the elderly—raged during the period leading up to the 2004 presidential and congressional elections. But let's focus on our first observation and question.

Observation 1: Health-care spending has increased faster than the rest of the U.S. economy since 1990.

How has health-care spending changed relative to the rest of the economy? To determine this, we need a measure of health-care spending and a measure of the size of the overall economy.

Spending as a Share of GDP

gross domestic product (GDP): a measure of the value of all the goods and services newly produced in an economy during a specified period of time.

The most comprehensive available measure of the size of an economy is the **gross domestic product (GDP).** For the United States, GDP is the total value of all products made in the United States during a specified period of time, such as a year. GDP includes all newly made goods, such as cars, trucks, shoes, airplanes, houses, and telephones; it also includes services, such as education, rock concerts, and health care. To measure the total value of all products made in the economy, economists add up the dollars that people spend on the products.

How large is GDP in the United States? In 2004, it was $11,735 billion, or about $11.7 trillion. We can compute GDP for any year. The question about health-care spending and the size of the economy requires that we look at the U.S. economy since 1990. In 1990, GDP was $5,803 billion. Column (1) of Table 2.1 provides a history of GDP since 1990.

Graphs are frequently a more helpful way to present data like those shown in Table 2.1. Figure 2.1 plots the data on GDP from column (1) of Table 2.1. The vertical axis is measured in billions of dollars; the horizontal axis is measured in years. For example, the point at the extreme lower left in Figure 2.1 represents GDP of $5,803 billion (on the vertical axis) in the year 1990 (on the horizontal axis). The points are connected by a line, which helps us visualize the steady growth of GDP during this period.

Now let us consider health-care spending, which includes payments for hospital services, lab tests, nursing homes, visits to the doctor or dentist, drugs, hearing aids, and eyeglasses. If we add up all spending on health care, we get $1,392 billion in 2004.

Table 2.1
GDP and Health-Care Spending, 1990–2004

Year	(1) GDP	(2) Health-Care Spending	(3) Health-Care Share of GDP (percent)
1990	5,803	556	9.6
1991	5,996	609	10.2
1992	6,338	672	10.6
1993	6,657	715	10.7
1994	7,073	753	10.6
1995	7,398	798	10.8
1996	7,817	833	10.7
1997	8,304	873	10.5
1998	8,747	921	10.5
1999	9,268	961	10.4
2000	9,817	1,027	10.5
2001	10,128	1,114	11.0
2002	10,487	1,210	11.5
2003	11,004	1,301	11.8
2004	11,735	1,392	11.9

Note: GDP and health-care spending are measured in billions of dollars.
Source: The source of all data in this section is the U.S. Department of Commerce National Income and Product Accounts.

100 times column 2 divided by column 1

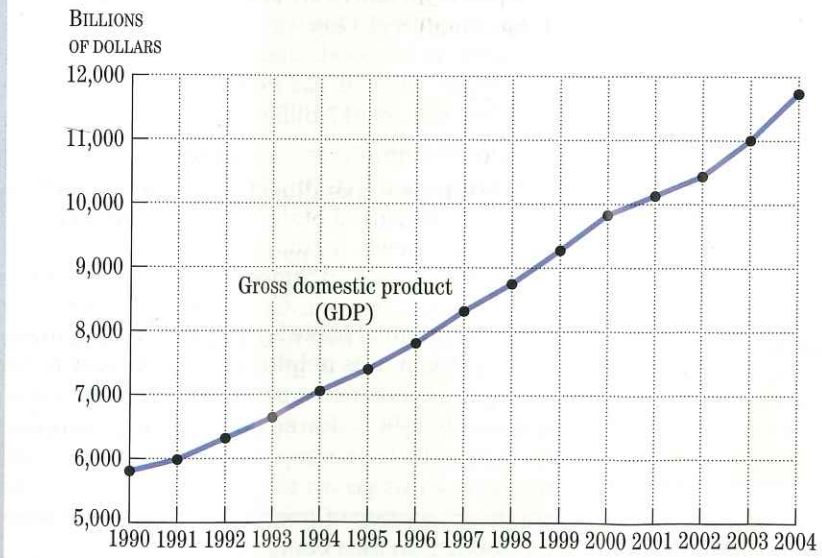


Figure 2.1
Gross Domestic Product (GDP) in the United States, 1990–2004
GDP is the total dollar value of newly produced goods and services. It can be measured by adding up what people spend on everything, from health care to cars. For each year from 1990 to 2004, GDP is plotted; the line connects all the points.

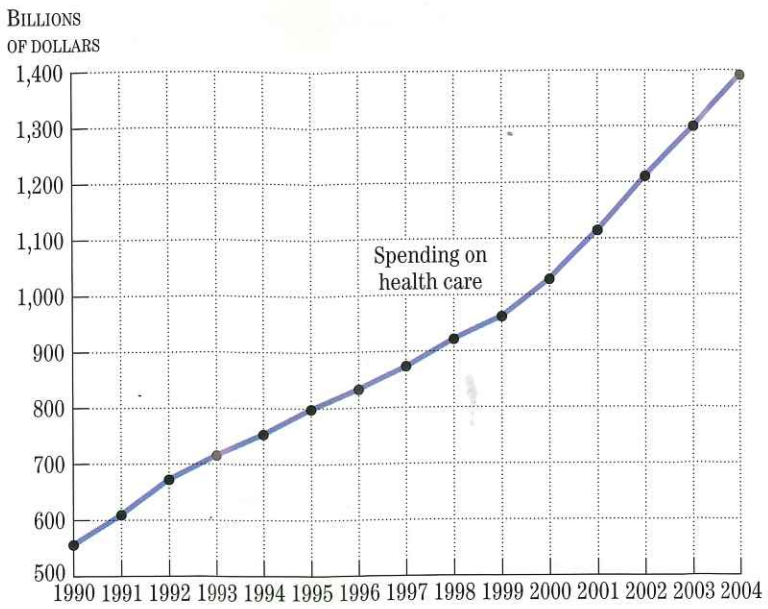


Figure 2.2
Spending on Health Care in the United States, 1990–2004
Health-care spending is the dollar value of payments for hospitals, doctors, dentists, nursing homes, drugs, and other items that provide medical care. Health care is one part of GDP.

This amount is more than three times as large as the entire automobile industry, including cars, trucks, and parts. Health care is the biggest industry in the United States.

Health-care spending since 1990 is listed in column (2) of Table 2.1 and is plotted in Figure 2.2. Figures 2.1 and 2.2 show that both GDP and health-care spending have grown since 1990. One way to assess the growth of health-care spending compared to spending on all goods and services is to look at health-care spending as a share, or percentage, of GDP. For example, in 2000, health-care spending was $1,027 billion, and GDP was $9,817 billion. Thus, the share (in percentage terms) of GDP going to health care was:

$$\frac{\text{Health-care spending}}{\text{GDP}} \times 100 = \text{health-care spending as a share of GDP}$$

$$\frac{1{,}027}{9{,}817} \times 100 = 10.5 \text{ percent}$$

Observation 2: The price of health care has risen compared with the price of other goods and services in the economy.

Column (3) of Table 2.1 gives the results of this calculation for all years from 1990 to 2004. Again, it is helpful to plot the shares, as in Figure 2.3. Observe how health-care spending rose as a percentage of GDP—or relative to the size of the economy—in the early 1990s, slowed down a little in the late 1990s, but then increased again in the 2000s. We have now quantified the observation about rising health-care spending. Now let us go on to consider the second question about health care and see whether the price of health care has risen compared with the price of other goods and services in the economy.

The Relative Price

relative price: the price of a particular good compared to the price of other things.

To see whether the price of health care has increased more or less rapidly than other things, we look at the relative price of health care. The **relative price** is a measure of health-care prices compared with the average prices of all goods and services, computed by dividing the health-care price by the average price of all goods and services:

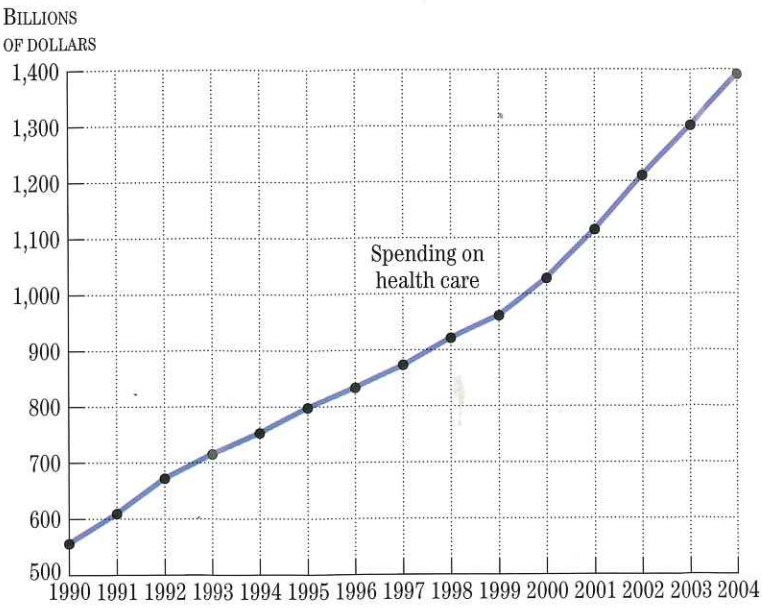


Figure 2.2
Spending on Health Care in the United States, 1990–2004
Health-care spending is the dollar value of payments for hospitals, doctors, dentists, nursing homes, drugs, and other items that provide medical care. Health care is one part of GDP.

This amount is more than three times as large as the entire automobile industry, including cars, trucks, and parts. Health care is the biggest industry in the United States.

Health-care spending since 1990 is listed in column (2) of Table 2.1 and is plotted in Figure 2.2. Figures 2.1 and 2.2 show that both GDP and health-care spending have grown since 1990. One way to assess the growth of health-care spending compared to spending on all goods and services is to look at health-care spending as a share, or percentage, of GDP. For example, in 2000, health-care spending was $1,027 billion, and GDP was $9,817 billion. Thus, the share (in percentage terms) of GDP going to health care was:

$$\frac{\text{Health-care spending}}{\text{GDP}} \times 100 = \text{health-care spending as a share of GDP}$$

$$\frac{1{,}027}{9{,}817} \times 100 = 10.5 \text{ percent}$$

Observation 2: The price of health care has risen compared with the price of other goods and services in the economy.

Column (3) of Table 2.1 gives the results of this calculation for all years from 1990 to 2004. Again, it is helpful to plot the shares, as in Figure 2.3. Observe how health-care spending rose as a percentage of GDP—or relative to the size of the economy—in the early 1990s, slowed down a little in the late 1990s, but then increased again in the 2000s. We have now quantified the observation about rising health-care spending. Now let us go on to consider the second question about health care and see whether the price of health care has risen compared with the price of other goods and services in the economy.

The Relative Price

relative price: the price of a particular good compared to the price of other things.

To see whether the price of health care has increased more or less rapidly than other things, we look at the relative price of health care. The **relative price** is a measure of health-care prices compared with the average prices of all goods and services, computed by dividing the health-care price by the average price of all goods and services:

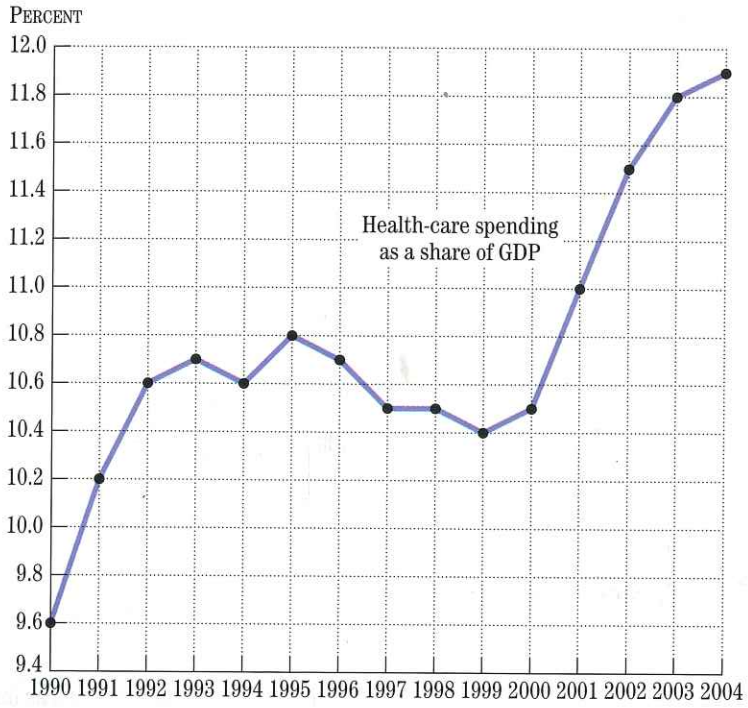


Figure 2.3
Health-Care Spending as a Share of GDP
Each point on the graph is the ratio of spending on health care from Figure 2.2 to GDP from Figure 2.1, expressed as a percentage. For example, in 2000, health-care spending was $1,027 billion and GDP was $9,817 billion. The ratio is 1,027/9,817 = .105, or 10.5 percent.

$$\text{Relative price of health care} = \frac{\text{health-care price}}{\text{average price of all goods and services}}$$

The relative price of health care is shown in Table 2.2 and plotted in Figure 2.4. Observe that the relative price of health care has risen since 1990. In other words, the price of health care has increased relative to the average of all other prices.

Correlations Between Economic Variables

economic variable: any economic measure that can vary over a range of values.

So far, we have focused on two economic variables to quantify our observations about health care in the United States. These variables are (1) health-care spending's share of GDP and (2) the relative price of health care. An **economic variable** is any economic measure that can vary over a range of values. Are these two economic variables correlated? Are there interesting patterns?

Figure 2.5 is useful for determining whether the relative price of health care and health care's share of GDP have been correlated. Each point in the figure corresponds to a relative price and a health-care share taken from Table 2.2 and from the last column of Table 2.1. The relative price is on the vertical axis, and health care as a share of GDP is on the horizontal axis.

The points in Figure 2.5 trace out some patterns. As the relative price of health care increased from 1990 to 1993, so did spending on health care as a share of GDP. Two variables are *correlated* if they tend to move up or down at the same time. There is a *positive correlation* if the two variables move in the same direction: When one goes up, the other goes up. Health-care spending as a share of GDP and the relative price of health care were positively correlated from 1990 to 1993 and from 2001 to 2004.

Table 2.2
Relative Price of Health Care

Year	Relative Health-Care Price
1990	0.887
1991	0.908
1992	0.936
1993	0.958
1994	0.977
1995	0.993
1996	1.000
1997	1.008
1998	1.022
1999	1.032
2000	1.039
2001	1.051
2002	1.060
2003	1.077
2004	1.087

Note: The relative price is a ratio of the price of health care to the average price of all goods and services. The ratio is set to 1 in 1996. This year is arbitrary: Using another year would not change the patterns of the relative prices.

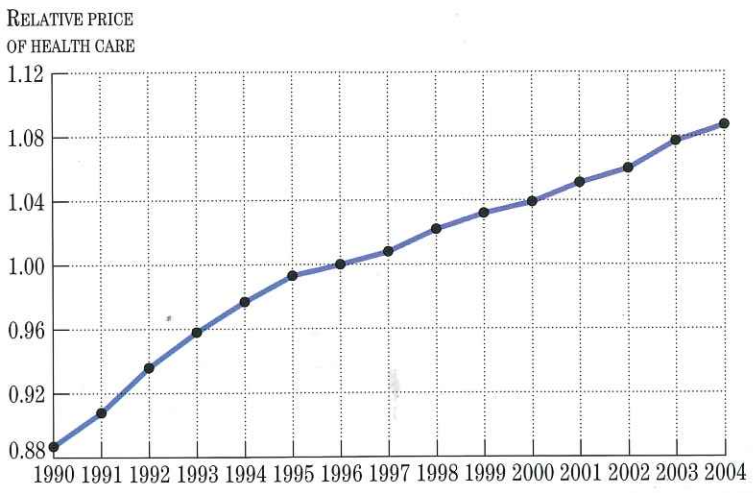


Figure 2.4
Relative Price of Health Care
Health-care prices rose more rapidly than other prices in the 1990s. Hence, the relative price of health care has increased; the increase was especially rapid in the early 1990s and the early 2000s.

Two variables are negatively correlated if they tend to move in opposite directions. From 1995 to 2000, a higher relative price of health care was associated with lower spending shares; thus, there was a *negative correlation* during those years.

■ **Correlation versus Causation.** Just because there is a correlation between two variables does not mean that one caused the other. There is a difference between *causation* and *correlation*. *Correlation* means that one event is usually observed to occur along with another. For example, high readings on a thermometer occur when it is hot outside. *Causation* means that one event brings about another event. But correlation does not imply causation. For example, the high reading on the thermometer does not cause the hot weather, even though the high reading and the hot weather are correlated. In this example, we know that the causation is the other way around: Hot weather causes the reading on the thermometer to be high.

More to the point for economics, if you looked only at the correlation in Figure 2.5 from 1990 to 1993, you might be tempted to say that the higher price of health care caused health-care spending to rise. But that correlation does not permit us to make such a conclusion about causation. We need to know more about the effects of the price of health care on health-care spending and about what determines the price before we can make statements about causality. In fact, between 1995 and 1999, spending on health care fell as a share of GDP as the relative price rose. One would have been proven wrong if one had argued that a higher price of health care would increase health-care spending.

controlled experiments: empirical tests of theories in a controlled setting in which particular effects can be isolated.

■ **The Lack of Controlled Experiments in Economics.** In many sciences—certainly psychology, medicine, and biology—investigators perform **controlled experiments** to determine whether one event causes another event. An example of a

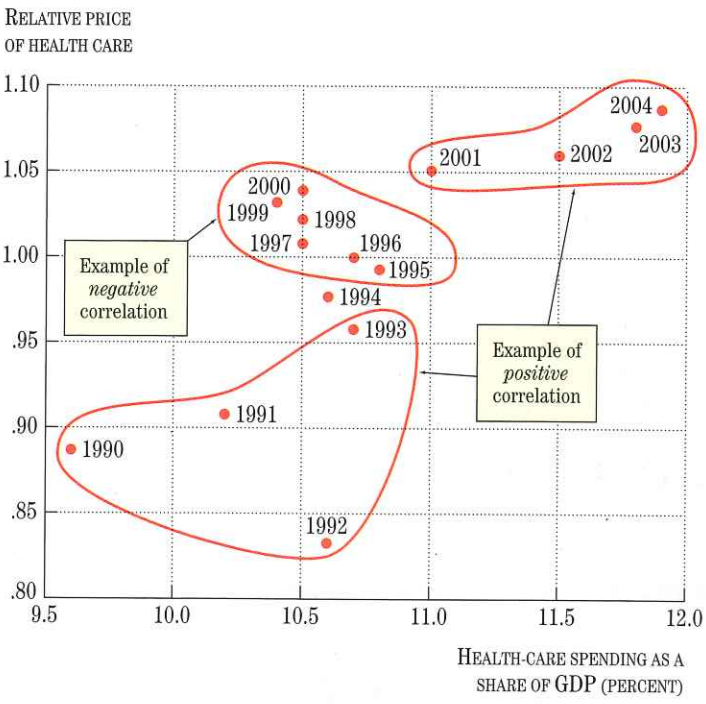


Figure 2.5
Relative Price of Health Care versus Health-Care Spending Share
The figure plots pairs of points: the relative price of health care on the vertical axis and health-care spending as a share of GDP on the horizontal axis. The observations come from Table 2.2 and the last column of Table 2.1.

controlled experiment is the clinical trial of new drugs. New drugs are tested by trying them out on two groups of individuals. One group gets the drug; the other group gets a placebo (a pill without the drug). If the experiment results in a significantly greater number of people being cured among the group taking the drug than among the control group not taking the drug, investigators conclude that the drug causes the cure.

Unfortunately, such controlled experiments are rare in economics. In the case of health-care prices and health-care spending, we cannot go back and repeat the years from 1990 to 2004 with a different health-care price and see what happens. True, we could look at other countries' experience, or we could look at the experience of different states within the United States. Economists use such comparisons to help determine causation. For example, we could look at one state in which the price of medical care increased and one in which it did not. We could then look at the health-care spending in each state to see if higher prices caused health-care spending to increase. But, unfortunately, no two countries or states are alike in all respects. Thus, attempting to control for other factors is not as easy as in the case of clinical trials.

In recent years, economists have adapted some methods of experimental science and have begun to conduct economic experiments in laboratory settings that are similar to the real world. The experiments can be repeated, and various effects can be controlled for. **Experimental economics** is a growing area of economics. The findings of experimental economics have affected economists' understanding of how the economy works. Experiments in economics also provide an excellent way to *learn* how the economy works, much as experiments in science courses can help one learn about gravity or the structure of plant cells. But because it is difficult to replicate real-world settings exactly in such experiments, they have not yet been applied as widely as the clinical or laboratory experiments in other sciences.

experimental economics: a branch of economics that uses laboratory experiments to analyze economic behavior.

■ **Faulty Data.** Economic data are not always accurate. People sometimes do not understand the survey questions, are too busy to fill them out carefully, or do not have the correct information. Hospitals reporting data on health-care prices, for example, may not take into account changes in the quality of health care. When people purchase medical care, the quality of the service provided can vary widely over time and from doctor to doctor.

If the quality of health care is improving, then the higher relative prices we have observed might partly reflect better service rather than an increase in the price of the same service. The actual increase in the price of health care might have been less rapid if we measured the improved quality, such as reduced chances of serious stroke or depression because of better drugs.

REVIEW

- Economists endeavor to explain facts and observations about the economy, but it is not always easy to establish what the facts are. To establish patterns, it is sometimes necessary to carefully organize information and present it in tables or graphs.
- GDP is a measure of all the goods and services produced in a country during a period of time.
- Correlation does not imply causation. Because controlled experiments are rare in economics, establishing causation is more difficult in economics than in other sciences.
- Recent advances in experimental economics are improving this situation.
- Economic observations are not always accurate. For example, the quality of a service such as medical care can be difficult to measure.

The Circular Flow Diagram: People Interacting in Markets

circular flow diagram: a diagram illustrating the flow of funds through the economy as people buy and sell in markets.

Behind the health-care and GDP observations in the figures and tables we've just examined are real people who purchase and produce health care and other services and goods in the economy. Any explanation of spending trends must be based on the choices these people make and how they interact with each other. In order to organize our thinking about who these people are and how they interact in the economy, a diagram is helpful. It is called the **circular flow diagram** and is shown in Figure 2.6. The circular flow diagram shows how funds flow through the economy as people buy and sell things. Notice the arrows pointing in a circular pattern (that's how the diagram gets its name). To understand how the diagram works, we need to discuss the boxes that the arrows point into and out of.

When using a circular flow diagram, economists place the things that people buy and sell into three groups: (1) *goods and services*, such as flu shots or medical physicals; (2) *labor*, such as the work of nurses and doctors; and (3) *capital*, such as the x-ray machines or the hospital beds that are needed to provide health care. In general, the term *capital* refers to the equipment and structures used to produce goods and services. Corresponding to each of these three groups is a market in which items in the group are bought and sold: (1) *the goods and services market*, (2) *the labor market*, and (3) *the capital market*. These three markets are shown by the blue boxes in the middle part of Figure 2.6, with the light blue boxes giving examples from the health-care industry.

Now consider the *households* and the *firms* in the circular flow diagram. A household is an individual or group of individuals occupying a set of living quarters. For example, a household could be a group of college graduates sharing an apartment, a divorced person living alone, a family with four children, or a single retiree. A firm is a producer of goods and services. For example, a firm in the health-care industry could be a local hospital or a large health maintenance organization. Of course, there are many other types of producers in the economy: General Motors produces cars, the University of Texas produces educational services, the U.S. government produces defense services, and so on. The circular flow diagram puts all producers into the "firm" category.

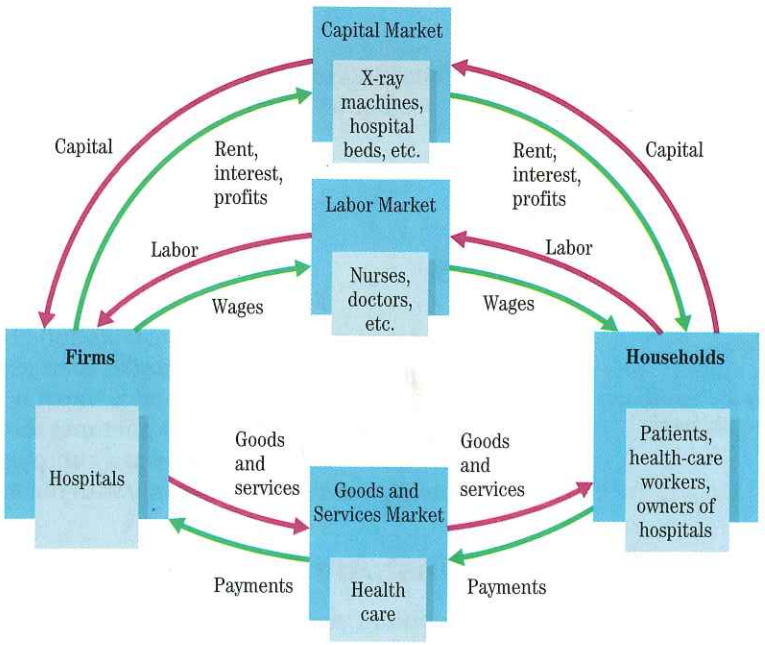


Figure 2.6
A Circular Flow Diagram
This diagram shows how funds flow through the economy from households to firms and back again. Buying and selling take place in the goods and services markets, the labor market, and the capital market.

The households and firms are shown in the blue boxes on the right and left, respectively, of the circular flow diagram, with the light blue boxes again giving examples from the health-care industry. The households are shown doing three things in the circular flow diagram:

- Households buy goods and services; for example, an older person gets a flu shot at the doctor's office.
- Households sell their labor services to firms in the labor market; for example, a nurse works for an HMO and receives wages in return.
- Households supply capital to firms in the capital market; for example, a young couple may rent out the first floor of their house as a doctor's office and receive rent in return. Households also supply capital by owning shares in firms (and receiving part of the profits) or making loans to firms (and receiving interest).

Note that in each of these cases, firms are on the other side of the market from households. Thus the firms are also shown doing three things: They sell goods and services to households, they buy labor services from households, and they buy capital from households.

The arrows in the circular flow diagram show the results of all this buying and selling. The counterclockwise (red) arrows show the movement of goods and services, labor, and capital. The clockwise (green) arrows show the flow of funds used to pay for these items: payments for goods and services, wages for labor, and rent, interest, and profit for capital. Funds flow through the economy from households to firms and back again.

The circular flow diagram is a useful visual device for keeping track of the people in the economy and the markets in which they interact. If we are going to understand the workings of any part of the economy—including health care—we must think

about these households and firms as they interact in the markets. That is where economic theory or models come into play.

REVIEW ▪ The circular flow diagram is useful for showing how households and firms interact in markets and how funds flow through the economy.

Economic Models

economic model: an explanation of how the economy or part of the economy works.

In order to explain economic facts and observations, one needs an economic theory, or a *model*. An **economic model** is an explanation of how the economy or a part of the economy works. In practice, most economists use the terms *theory* and *model* interchangeably, although sometimes the term *theory* suggests a general explanation and the term *model* suggests a more specific explanation. The term *law* is also typically used interchangeably with the terms *model* and *theory* in economics.

What Are Economic Models?

Economic models are always abstractions, or simplifications, of the real world. They take very complicated phenomena, such as the behavior of people, firms, and governments, and simplify them, in much the same way as a model of a building used by architects is an abstraction, or simplification, of the actual building. Some models can be very detailed; others are just broad abstractions. Be sure to remember that the model and the phenomenon being explained by the model are different.

Do not be critical of economic models just because they are simplifications. In every science, models are simplifications of reality. Models are successful if they explain reality reasonably well. In fact, if they were not simplifications, they would be hard to use effectively. Economic models differ from those in the physical sciences because they endeavor to explain human behavior, which is complex and often unpredictable. It is for this reason that the brilliant physicist Max Planck said that economics was harder than physics.

Economic models can be described with words, with numerical tables, with graphs, or with algebra. To use economics, it is important to be able to work with these different descriptions. Figures 2.7 and 2.8 show how models can be illustrated with graphs. By looking at graphs, we can see quickly whether the model has an inverse or a direct relationship. If a model says that one variable varies inversely with the other, this means that if the first variable rises, then the second falls. If a model says that one variable varies directly with another, this means that if one variable rises, the other also rises. In economics, the expression "is positively related to" is frequently used in place of the phrase "varies directly with," which is more common in other sciences. Similarly, the expression "is negatively related to" is frequently used in place of "varies inversely with."

positively related: a situation in which an increase in one variable is associated with an increase in another variable; also called *directly related.*

negatively related: a situation in which an increase in one variable is associated with a decrease in another variable; also called *inversely related.*

In Figure 2.7, two variables—perhaps a relative price variable and a spending variable—are shown to be **positively related.** In other words, when variable 1 increases from *A* to *B*, variable 2 increases from *C* to *D* by the specific amount given by the curve. Likewise, when variable 1 decreases from *B* to *A*, variable 2 decreases from *D* to *C*. In Figure 2.8, a model with two variables that are **negatively related** is shown. Here, when variable 1 increases from *A* to *B*, variable 2 decreases from *D* to *C*. Likewise, when variable 1 decreases from *B* to *A*, variable 2 increases from *C* to *D*. Models have *constants* as well as variables. The constants in the models in Figures 2.7 and 2.8 are the positions and shapes of the curves.

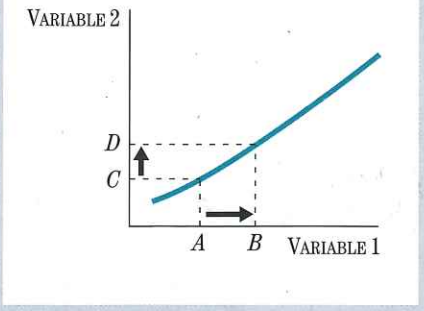


Figure 2.7
A Model with Two Positively Related Variables
The upward-sloping line shows how the variables are related. When one variable increases from *A* to *B*, the other variable increases from *C* to *D*. If one variable declines from *B* to *A*, the other variable declines from *D* to *C*. We say that variable 1 is positively related to variable 2, or that variable 1 varies directly with variable 2.

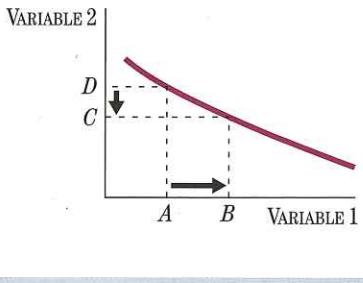


Figure 2.8
A Model with Two Negatively Related Variables
When one variable increases from *A* to *B*, the other variable decreases from *D* to *C*. Likewise, when one variable decreases from *B* to *A*, the other variable increases from *C* to *D*. We say that variable 1 is negatively related to variable 2, or that variable 1 varies inversely with variable 2.

As you study economic models in this book, you will begin to see that they have a common approach to human behavior. Economic models describe how people deal with scarcity, or with situations where they would like to have more, or do more, than their resources will allow. Scarcity requires that people *choose* between one thing and another. For example, in developing a model of health-care spending, an economist would examine how people with a limited amount of income would choose between more health care and less of something else. Or the economist might examine how a health maintenance organization would decide whether to hire more nurses and fewer doctors in producing a given amount of health care in the community. In modeling the behavior of consumers and firms, economists assume that people make purposeful choices to improve their well-being when confronting such scarcity. More than anything else, the problem of scarcity provides a broad and common core to the field of economics.

■ **An Example: A Model with Two Variables.** Figure 2.9 shows a model describing how doctors employed in a health maintenance organization provide physical examinations. The model states that the more doctors who are employed at the HMO, the more physical exams can be given. The model is represented in four different ways: (1) with words, (2) with a numerical table, (3) with a graph, and (4) with algebra.

On the lower right of Figure 2.9, we have the verbal description: more doctors, more physical exams, but additional doctors increase the number of exams by smaller amounts, presumably because the diagnostic facilities at the HMO are limited; for example, there are only so many rooms available for physical exams.

On the upper left, we have a table with numbers showing how the number of examinations depends on the number of doctors. Exactly how many examinations are given by each number of doctors is shown in the table. Clearly this table is much more specific than the verbal description. Be sure to distinguish between the meaning of a table that presents a model (like the table in Figure 2.9) and a table that

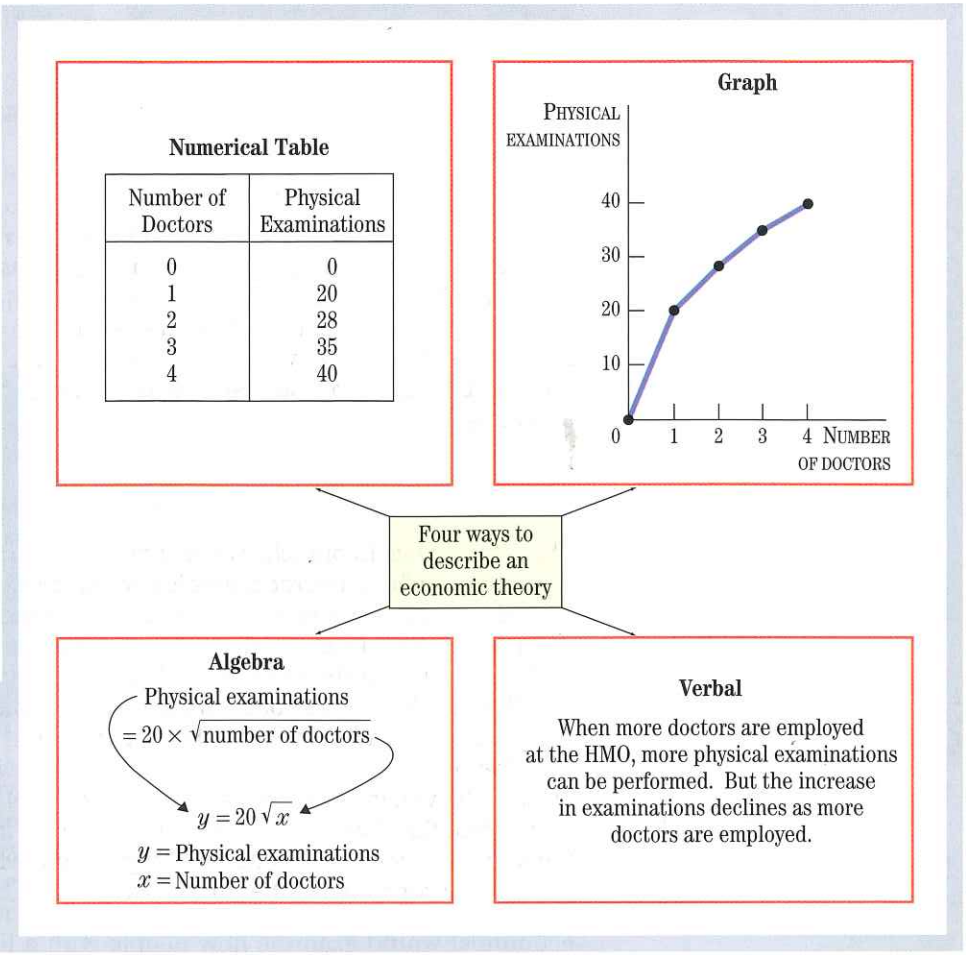


Figure 2.9
Economic Models in Four Ways
Each way has advantages and disadvantages; this book focuses mostly on verbal descriptions, graphs, and numerical tables, but occasionally some algebra will be used to help explain things.

presents data (like Table 2.1). They look similar, but one is a model of the real world and the other represents observations about the real world.

On the upper right, we have a curve showing the relationship between doctors and physical examinations. The curve shows how many exams each number of doctors can perform. The points on the curve are plotted from the information in the table. The vertical axis has the number of examinations; the horizontal axis has the number of doctors. The points are connected with a line to help visualize the curve.

Finally, in the lower left we show the doctor-examination relationship in algebraic form. In this case, the number of exams is equal to the square root of the number of doctors times 20. If we use the symbol y for the number of exams and x for the number of doctors, the model looks a lot like the equations in an algebra course.

All four ways of representing models have advantages and disadvantages. The advantage of the verbal representation is that we usually communicate with people in words, and if we want our economic models to have any use, we need to communicate with people who have not studied economics. However, the verbal representation is not as precise as the other three. In addition to verbal analysis, in this book we will focus on tabular and graphical representations rather than on algebraic descriptions.

■ **Prediction and the *Ceteris Paribus* Assumption.** Prediction is one of the most important uses of models. For example, using the model in Figure 2.8, we can

predict that if variable 1 rises from *A* to *B*, then variable 2 will fall from *D* to *C*. Using the model for physical exams at an HMO, we might predict that having more doctors at the HMO will increase the number of physicals that can be given. Economists use models to predict variables ranging from GDP next year to the price of medical care in the year 2007.

***ceteris paribus*:** "all other things being equal"; refers to holding all other variables constant or keeping all other things the same when one variable is changed.

In order to use models for prediction, economists use the assumption of ***ceteris paribus*,** which means "all other things being equal." For example, the prediction that variable 2 will fall from *D* to *C* assumes that the curve in Figure 2.8 does not shift: The position of the curve when variable 1 is at *A* is *equal* to the position of the curve when variable 1 is at *B*. If other things were not equal—if the curve shifted—then we could not predict that variable 2 would fall from *D* to *C* when variable 1 rose from *A* to *B*. Similarly, predicting that more doctors can produce more physical exams assumes that there is no power outage that would cause the diagnostic equipment to stop operating.

Microeconomic Models versus Macroeconomic Models

There are two main branches of economics: microeconomics and macroeconomics; thus, there are both microeconomic and macroeconomic models.

microeconomics: the branch of economics that examines individual decision-making at firms and households and the way they interact in specific industries and markets.

Microeconomics studies the behavior of individual firms and households or specific markets like the health-care market or the college graduate market. It looks at variables such as the price of a college education or the reason for increased wages of college graduates. Microeconomic models explain why the price of gasoline varies from station to station and why there are discount airfares.

macroeconomics: the branch of economics that examines the workings and problems of the economy as a whole—GDP growth and unemployment.

Macroeconomics focuses on the whole economy—the whole national economy or even the whole world economy. It tries to explain the changes in GDP over time rather than the changes in a part of GDP like health-care spending. It looks at questions such as what causes the GDP to grow and why many more workers are unemployed in Europe than in the United States.

The Use of Existing Models

Because economics has been around for a long time, there are many existing models that can be applied to explain observations or make predictions that are useful to decision-makers. Much of what economists do in practice, whether in government or business or universities, is use models that are already in existence.

The models are used in many different types of applications, from determining the effects of discrimination in the workplace to evaluating the gains from lower health-care prices. Frequently the models are applied in new and clever ways.

The Development of New Models

Like models in other sciences, economic models change and new models are developed. Many of the models in this book are very different from the models in books published 40 years ago. New economic models evolve because some new observations cannot be explained by existing models.

The process of the development of new models or theories in economics proceeds much like that in any other science. First one develops a *hypothesis*, or a hunch, to explain a puzzling observation. Then one tests the hypothesis by seeing if its predictions of other observations are good. If the hypothesis passes this test, then it becomes accepted. In practice, however, this is at best a rough description of the process of scientific discovery in economics. Existing models are constantly being

reexamined and tested. Some economists specialize in testing models; others specialize in developing them. There is an ongoing process of creating and testing of models in economics.

REVIEW

- Economists use economic models to explain economic observations. Economic models are similar to models in other sciences. They are abstractions, or simplifications, of reality, and they have variables and constants. But economic models are different from models in the physical sciences because they must deal with human behavior. Models can be represented verbally, with numerical tables, with graphs, and with algebra.
- New economic models are developed in part because existing models cannot explain facts or observations.

The Impact of Economics on Public Policy

Ever since the birth of economics as a field—around 1776, when Adam Smith published the *Wealth of Nations*—economists have been concerned about and motivated by a desire to improve the economic policy of governments. In fact, economics was originally called *political economy*. Much of the *Wealth of Nations* is about what the government should or should not do to affect the domestic and international economy.

Adam Smith argued for a system of *laissez faire*—little government control—where the role of the government is mainly to promote competition, provide for national defense, and reduce restrictions on the exchange of goods and services. One hundred years later, Karl Marx brought a new perspective to Smith's (and other classical economists') view of political economy, arguing against the laissez-faire approach. His analysis of market economies, or **capitalism,** centered on the contradictions that he saw arising out of such a system, particularly the conflict between the owners of production and laborers. He argued that these contradictions would result in the inevitable collapse of capitalism and the emergence of a new economic system, called **socialism,** in which government would essentially own and control all production. While Marx actually wrote very little about what a socialist or communist economy would look like, the centrally planned economies that arose in the Soviet Union, Eastern Europe, and China in the twentieth century can be traced to Marx's ideas. Most countries today have rejected the command economy and have moved toward market economies, but the debate about the role of government continues. In many modern market economies, the government plays a large role, and for this reason, such economies are sometimes called **mixed economies.** How great should the role of government be in a market economy? Should the government provide health-care services? Should it try to break up large firms?

capitalism an economic system based on a market economy in which capital is individually owned, and production and employment decisions are decentralized.

socialism an economic system in which the government owns and controls all the capital and makes decisions about prices and quantities as part of a central plan.

mixed economy: a market economy in which the government plays a very large role.

Positive versus Normative Economics

In debating the role of government in the economy, economists distinguish between positive and normative economics. **Positive economics** is about what is; **normative economics** is about what should be. For example, positive economics endeavors to explain why health-care spending slowed down in the mid-1990s.

positive economics: economic analysis that explains what happens in the economy and why, without making recommendations about economic policy.

normative economics: economic analysis that makes recommendations about economic policy.

Council of Economic Advisers: a three-member group of economists appointed by the president of the United States to analyze the economy and make recommendations about economic policy.

Normative economics aims to develop and recommend policies that might prevent health-care spending from rising rapidly in the future. In general, normative economics is concerned with making recommendations about what the government should do—whether it should control the price of electricity or health care, for example. Economists who advise governments spend much of their time doing normative economics. In the United States, the president's **Council of Economic Advisers** has legal responsibility for advising the president about which economic policies are good and which are bad.

Positive economics can also be used to explain *why* governments do what they do. For example, why did the U.S. government control airfares and then stop? Why were tax rates cut in the 1980s, increased in the 1990s, then cut again in the 2000s? Positive analysis of government policy requires a mixture of both political science and economic science, with a focus on what motivates voters and the politicians they elect.

Economics as a Science versus a Partisan Policy Tool

Although economics, like any other science, is based on facts and theories, it is not always used in a purely scientific way.

In political campaigns, economists put forth arguments in favor of one candidate, emphasizing the good side of their candidate and de-emphasizing the bad side. In a court of law, one economist may help a defendant—making the best case possible—and another economist may help the plaintiff—again, making the best case possible. In other words, economics is not always used objectively. A good reason to learn economics for yourself is to see through fallacious arguments.

But economics is not the only science that is used in these two entirely different modes. For example, there is currently a great controversy about the use of biology and chemistry to make estimates of the costs and benefits of different environmental policies. This is a politically controversial subject, and some on both sides of the controversy have been accused of using science in nonobjective ways.

Economics Is Not the Only Factor in Policy Issues

Although economics can be very useful in policy decisions, it is frequently not the only factor. For example, national security sometimes calls for a recommendation on a policy issue different from one based on a purely economic point of view. Although most economists recommend free exchange of goods between countries, the U.S. government restricted exports of high-technology goods such as computers during the cold war because defense specialists worried that the technology could help the military in the Soviet Union, and this was viewed as more important than the economic argument. There are still heavy restrictions on trade in nuclear fuels for fear of the proliferation of nuclear weapons.

Disagreement Between Economists

Watching economists debate issues on television or reading their opinions in a newspaper or magazine certainly gives the impression that they rarely agree. There are major controversies in economics, and we will examine them in this book. But when people survey economists' beliefs, they find a surprising amount of agreement.

Why, then, the popular impression of disagreement? Because there are many economists, and one can always find some economist with a different viewpoint. When people sue other people in court and economics is an issue, it is always

Science or Persuasion?

In a recent court case, a grocery store chain, Lucky Stores, was sued for discriminating against female workers. The case illustrates how economics can be used in a partisan as well as a scientific way.

Economists were called as expert witnesses for both sides. Labor economist John Pencavel testified for the plaintiffs, the women who brought the suit. He found that women at Lucky earned between 76 and 82 percent of what Lucky's male workers earned. Pencavel found that women were regularly placed in jobs that paid less than jobs given male coworkers, although there was no significant difference between them in terms of education and experience. There was little difference in the wages of the male and female workers within each type of job, but some jobs paid more than others, and women happened to be assigned to the lower-paying jobs.

Joan Haworth, another labor economist, was an expert witness for the defendant, Lucky Stores. She reported survey evidence showing that Lucky's assignment of women and men to different jobs reflected differences in the work preferences of men and women. Thus, Lucky justified its job assignments by arguing that there was a gender difference in attitudes toward work. Lucky argued that its employment policies were based on observed differences in the career aspirations of male and female employees. For example, one manager at Lucky testified that women were more interested in cash register work and men were more interested in floor work.

After weighing the facts and economic arguments, the judge decided the case in favor of the plaintiffs.

You be the judge. Would you have been persuaded by the economic argument used by Lucky Stores or by the defendants?

Although male and female employees received equal pay for equal work, the judge concluded that Lucky's employment policies involved discrimination. The judge wrote: "The court finds defendant's explanation that the statistical disparities between men and women at Lucky are caused by differences in the work interests of men and women to be unpersuasive."

The decision is a landmark because of the economic analysis that showed that discrimination could exist even if men and women were being paid the same wage for equal work. Of course, not all sex discrimination cases are decided in favor of the plaintiffs. But whoever wins a given case, economics is almost always a key consideration in the judge's decision.

possible to find economists who will testify for each side, even if 99 percent of economists would agree with one side. Similarly, television interviews or news shows want to give both sides of public policy issues. Thus, even if 99 percent of economists agree with one side, it is possible to find at least one on the other side.

Economists are human beings with varying moral beliefs and different backgrounds and political views that are frequently unrelated to economic models. For example, an economist who is very concerned about the importation of drugs into the United States might appear to be more willing to condone a restriction on coffee exports from Brazil and other coffee-exporting countries, which might give Colombia a higher price for its coffee to offset a loss in revenue from cocaine. Another economist, who felt less strongly about drug imports, might argue strongly against such a restriction on coffee. But if they were asked about restrictions on trade in the abstract, both economists would probably argue for government policies that prevent them.

REVIEW

- Economic theory can be used to make better economic decisions. Improving government policy decisions has long been a purpose of economics.
- The most basic economic policy questions concern the general role of government in a market economy.

Conclusion: A Reader's Guide

In Chapter 1, we explored the central idea of economics: scarcity, choice, and economic interaction. In this chapter, we discussed how economists observe economic phenomena and use economic models to explain these phenomena. It is now time to move on and learn more about the models and application of the central idea. As you study economic models in the following chapters, it will be useful to keep three points in mind. They are implied by the ideas raised in this chapter.

First, *economics—more than other subjects—requires a mixture of verbal and quantitative skills.* Frequently, those who come to economics with a good background in physical science and mathematics find the mix of formal models with more informal verbal descriptions of markets and institutions unusual and perhaps a little difficult. If you are one of these people, you might wish for a more cut-and-dried, or mathematical, approach.

In contrast, those who are good at history or philosophy may find the emphasis on formal models and graphs difficult and might even prefer a more historical approach that looked more at watershed events and famous individuals and less at formal models of how many individuals behave. If you are one of these people, you might wish that economic models were less abstract.

In reality, however, economics is a mixture of formal modeling, historical analysis, and philosophy. If you are very good at math and you think the symbols and graphs of elementary economics are too simple, think of Max Planck's comment about economics and focus on the complexity of the economic phenomena that these simple models and graphs are explaining. Then when you are asked an open-ended question about government policy that does not have a simple yes or no answer, you will not be caught off guard. Or if your advantage is in history or philosophy, you should spend more time honing your skills at using models and graphs. Then when you are asked to solve a cut-and-dried economic problem with an exact answer requiring graphical analysis, you will not be caught off guard.

Second, *economics is about more than the stock market.* When your friends or relatives hear that you are taking economics, they may ask you for advice about what stock to buy. Economists' friends and relatives are always asking them for such advice. But economics alone offers no predictions about the success of particular companies. Rather, economics gives you some tools you can use to obtain information about companies and to analyze them yourself—perhaps eventually to become an investment adviser.

Economics will also help you answer questions about whether to invest in the stock market or in a bank or how many stocks to buy. But the scope of economics is much, much broader than the stock market or banks, as the questions at the start of the chapter indicate. In fact, the scope of economics is even wider than these

examples. Economists use their models, or theories, to study environmental pollution, crime, discrimination, and who should have the right to sue whom.

Third, and perhaps most important, *the study of economics is an intellectually fascinating adventure in its own right.* Yes, economics is highly relevant, and it affects people's lives. But once you learn how economic models work, you will find that they are actually fun to use. And they would be just as much fun if they were not so relevant. Every now and then, just after you have learned about a new economic model, put the book down and think of the economic model independent of its message or relevance to society—try to enjoy it the way you would a good movie. In this way, too, you will be learning to think like an economist.

KEY POINTS

1. Economics is a way of thinking that requires observation (describing economic events), building and using economic models to explain economic events and predict future events, and recommending courses of action for government—and business—based on these observations and models.
2. Economic models are abstractions, or simplifications, of reality and attempt to explain human behavior as expressed by economic measures.
3. Economic models, like models in other sciences, can be described with words, with tables, with graphs, or with mathematics. All four ways are important and complement one another.
4. The circular flow diagram shows the major players in the economy and how they interact in markets.
5. A plot showing that two variables are correlated during a period of time does not mean that one causes the other.
6. Faulty data and the lack of controlled experiments sometimes make economic observations difficult to interpret.
7. Sometimes new facts require that economists develop new models.
8. Decisions about the role of government in areas from airfares to health care are influenced by economic analysis.
9. Improving economic policy has been a goal of economists since the time of Adam Smith.

KEY TERMS

gross domestic product (GDP)
relative price
economic variable
controlled experiments
experimental economics
circular flow diagram
economic model
positively related
negatively related
ceteris paribus
microeconomics
macroeconomics
capitalism
socialism
mixed economy
positive economics
normative economics
Council of Economic Advisers

QUESTIONS FOR REVIEW

1. Why do economists need to document and quantify observations about the economy?
2. What is the most comprehensive available measure of the size of an economy?
3. What is meant by a relative price?
4. Why doesn't correlation imply causation?
5. What do the arrows in the circular flow diagram indicate?
6. How do economic models differ from the economic phenomena they explain?
7. Why are controlled experiments rare in economics?
8. How do economists use the *ceteris paribus* assumption?
9. What is the difference between macroeconomics and microeconomics?
10. What is the difference between positive and normative economics?

PROBLEMS

1. Which of the following items are microeconomic, and which are macroeconomic?
   a. The number of people with jobs in the United States
   b. A tax on sport utility vehicles
   c. Prices of sunglasses
   d. GDP

2. Identify whether the following policy statements are positive or normative. Explain.
   a. "The price of Internet stocks is too high."
   b. "The government should control the price of health care."
   c. "Increases in consumer spending improved the economy last year."
   d. "The government should break up Microsoft."
3. Interpret the data on spending on clothing in the table below by filling in the blanks.
   a. What has happened to clothing spending as a share of GDP over this 30-year period?
   b. What has happened to the relative price of clothing over this period, and how could it be related to the clothing spending share?
   c. Draw a graph showing the relationship between clothing spending as a share of GDP and the relative price of clothing.

Year	GDP (billions of dollars)	Spending on Clothing	Clothing Spending as a Share of GDP (%)	Relative Clothing Price
1970	1,039.7	47.8		1.88
1980	2,795.6		3.84	1.43
1990		204.1	3.52	1.17
2000	9,817.0	319.1		0.89

4. Draw a diagram like Figure 2.6 for the market for air travel and give examples of capital, labor, and firms in that market. Show that there is a complete circular flow of funds and goods in the economy as a whole.
5. Why is it typical for economists to make the *ceteris paribus* assumption when making predictions?
6. Consider an economic model of web page production. Show how to represent this model graphically, algebraically, and verbally, as in Figure 2.9.

Number of Programmers	Web Pages
0	0
1	10
4	20
9	30
16	40

7. What is the difference between the price of a good and its relative price? Which information is more useful if you are interested in analyzing the change in spending on that good?
8. Indicate whether you expect positive or negative correlation for the following pairs of variables. What is required in order to show causation?
   a. Sunrise and crowing roosters
   b. Price of theater tickets and number of theatergoers
   c. Purchases of candy and purchases of Valentine's Day cards

Reading, Understanding, and Creating Graphs

Whether you follow the stock market, the health-care market, or the whole economy, graphs are needed to understand what is going on. That is why the financial pages of newspapers contain so many graphs. Knowing how to read, understand, and even create your own graphs is part of learning to "think like an economist." Graphs help us see correlations, or patterns in economic observations. Graphs are also useful for understanding economic models. They help us see how variables in the model behave. They help us describe assumptions about what firms and consumers do.

Computer software to create graphs is now widely available. A graphing program with many examples is provided with the software that accompanies this text. To understand how helpful graphs can be, you might want to create a few of your own graphs using the time-series data in the "Explore" section of the software. Here we provide an overview of basic graphing techniques.

Visualizing Observations with Graphs

Most economic graphs are drawn in two dimensions, like the surface of this page, and are constructed using a **Cartesian coordinate system.** The idea of Cartesian coordinates is that pairs of observations on variables can be represented in a plane by designating one axis for one variable and the other axis for the other variable. Each point, or coordinate, on the plane corresponds to a pair of observations.

Time-Series Graphs

In many instances, we want to see how a variable changes over time. Consider the federal debt held by the public—all the outstanding borrowing of the federal government that has not yet been paid back. Table 2A.1 shows observations of the U.S. federal debt. The observations are for every 10 years. The observations in Table 2A.1 are graphed in Figure 2A.1. The graph in Figure 2A.1 is called a **time-series graph** because it plots a series—that is, several values of the variable—over time.

Observe the scales on the horizontal and vertical axes in Figure 2A.1. The seven years are put on the horizontal axis, spread evenly from the year 1950 to the year 2010. The last year is a forecast. For the vertical axis, one needs to decide on a scale. The range of variation for the debt in Table 2A.1 is very wide—from a minimum of $219 billion

Table 2A.1
U.S. Federal Government Debt

Year	Debt (billions of dollars)
1950	219
1960	237
1970	283
1980	712
1990	2,412
2000	3,410
2010 (Projected)	5,949

Source: Congressional Budget Office.

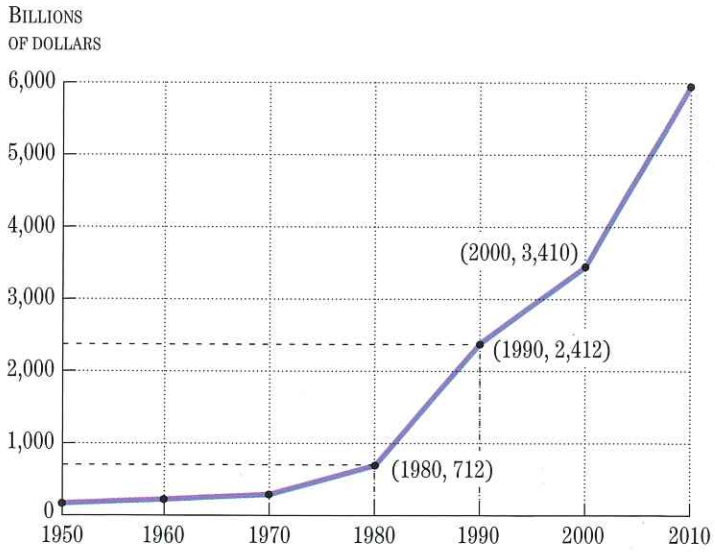
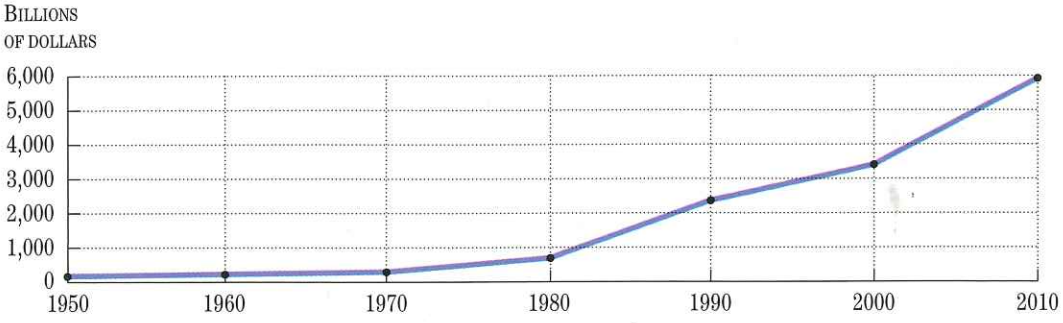
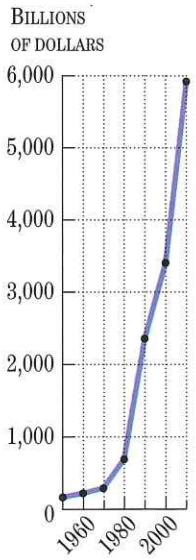


Figure 2A.1
U.S. Federal Debt
Each point corresponds to a pair of observations—the year and the debt—from Table 2A.1.

Figure 2A.2
Stretching the Debt Story in Two Ways
The points in both graphs are identical to those in Figure 2A.1, but by stretching or shrinking the scales the problem can be made to look either less dramatic or more dramatic.





to a maximum of $5,949 billion. Thus, the range on the vertical axis—from $0 to $6,000 billion in Figure 2A.1—must be wide enough to contain all these points.

Now observe how each pair of points from Table 2A.1 is plotted in Figure 2A.1. The point for the pair of observations for the year 1950 and the debt of $219 billion is found by going over to 1950 on the horizontal axis, then going up to $219 billion and putting a dot there. The point for 1960 and $237 billion and all the other points are found in the same way. In order to better visualize the points, they can be connected with lines. These lines are not part of the observations; they are only a convenience to help in eyeballing the observations. The points for 1980, 1990, and 2000 are labeled with the pairs of observations corresponding to Table 2A.1, but in general there is no need to put in such labels.

One could choose scales different from those in Figure 2A.1, and if you plotted your own graph from Table 2A.1 without looking at Figure 2A.1, your scales would probably be different. The scales determine how much movement there is in a time-series graph. For example, Figure 2A.2 shows two ways to stretch the scales to make the increase and decrease in the debt look more or less dramatic. So as not to be fooled by graphs, therefore, it is important to look at the scales and think about what they mean.

As an alternative to time-series graphs with dots connected by a line, the observations can be shown on a bar graph, as in Figure 2A.3. Some people prefer the visual look of a bar graph, but, as is clear from a comparison of Figures 2A.1 and 2A.3, they provide the same information.

The debt as a percentage of GDP is given in Table 2A.2 and graphed in Figure 2A.4. Note that this figure makes the debt look very different from the way it looks in the first one. As a percentage of GDP, the debt fell from the end of World War II (when it was very large because of the war debt) until around 1980. It increased during the 1980s and declined in the 1990s, but has started to increase again in the 2000s.

Sometimes the data to be graphed have no observations close to 0, in which case including 0 on the vertical

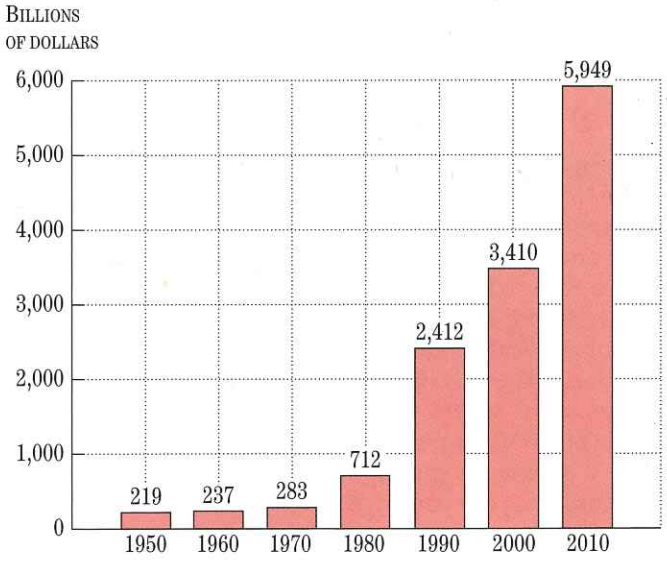


Figure 2A.3
U.S. Federal Debt in Bars
The observations are identical to those in Figure 2A.1.

Table 2A.2
U.S. Federal Debt as a Percentage of GDP

Year	Debt (percent of GDP)
1950	82.5
1960	46.8
1970	28.2
1980	25.7
1990	42.3
2000	35.9
2010 (Projected)	37.8

Source: U.S. Department of Commerce and Table 2A.1.

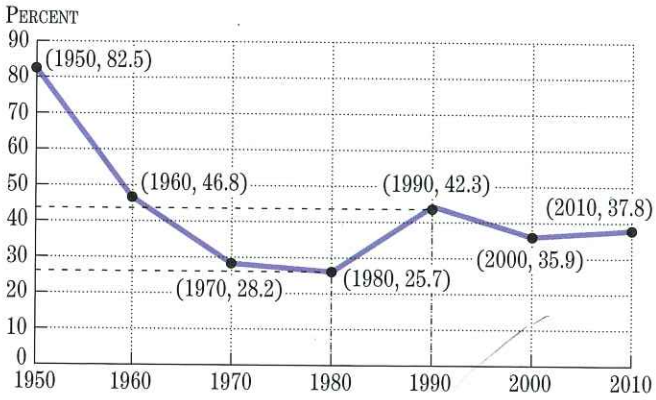


Figure 2A.4
U.S. Federal Debt as a Percentage of GDP
Each point corresponds to a pair of observations from Table 2A.2.

axis would leave some wasted space at the bottom of the graph. To eliminate this space and have more room to see the graph itself, we can start the range near the minimum value and end it near the maximum value. This is done in Figure 2A.5, where the debt as a percentage of GDP is shown *up to 1980*. Note, however, that cutting off the bottom of the scale could be misleading to people who do not look at the axis. In particular, 0 percent is no longer at the point where the horizontal and vertical axes intersect. To warn people about the missing part of the scale, a little cut is sometimes put on the axis, as is done in Figure 2A.5, but you have to look carefully at the scale.

Time-Series Graphs Showing Two or More Variables

So far, we have shown how a graph can be used to show observations on one variable over time. What if we want to see how two or more variables change over time

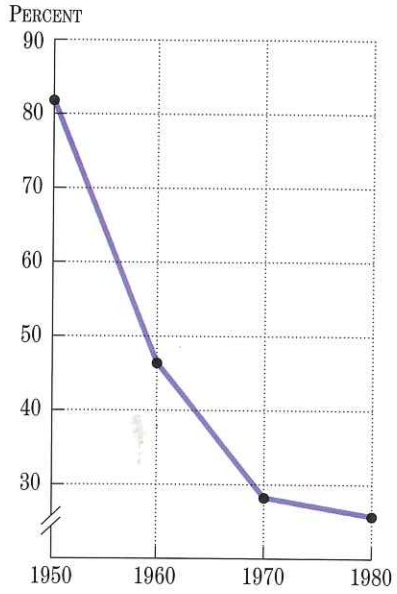


Figure 2A.5
A Look at Debt as a Percentage of GDP from 1950 to 1980
(*Note:* To alert the reader that the bottom part of the axis is not shown, a break point is sometimes used, as shown here.)

together? Suppose, for example, we want to look at how observations on debt as a percentage of GDP compare with the interest rate the government must pay on its debt. (The interest rate for 2010 is, of course, a forecast.) The two variables are shown in Table 2A.3.

The two sets of observations can easily be placed on the same time-series graph. In other words, we can plot the observations on the debt percentage and connect the dots and then plot the interest rate observations and connect the dots. If the scales of measurement of the two

Table 2A.3
Interest Rate and Federal Debt as a Percentage of GDP

Year	Debt (percent of GDP)	Interest Rate (percent)
1950	82.5	1.2
1960	46.8	2.9
1970	28.2	6.5
1980	25.7	11.5
1990	42.3	7.5
2000	35.9	5.5
2010 (Projected)	37.8	5.5

Source: Federal Reserve Board and Table 2A.2.

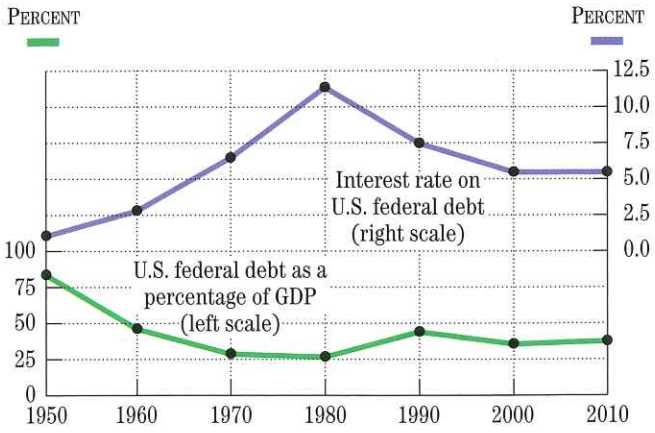


Figure 2A.6
Comparing Two Time Series with a Dual Scale
When two variables have different scales, a dual scale is useful. Here the interest rate and the debt as a percentage of GDP are plotted from Table 2A.3.

variables are much different, then it may be hard to see both, however. For example, the interest rate ranges between 1 and 12 percent; it would not be very visible on a graph going all the way from 0 to 100 percent, a range that is fine for the debt percentage. In this situation, a **dual scale** can be used, as shown in Figure 2A.6. One scale is put on the left-hand vertical axis, and the other scale is put on the right-hand vertical axis. With a dual-scale diagram, it is very important to be aware of the two scales. In Figure 2A.6 we emphasize the different axes by the color line segment at the top of each vertical axis. The color line segment corresponds to the color of the curve plotted using that scale.

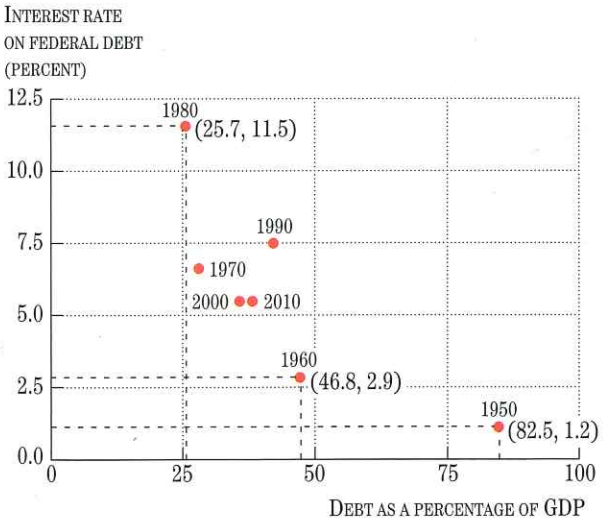


Figure 2A.7
Scatter Plot
Interest rate and debt as a percentage of GDP are shown.

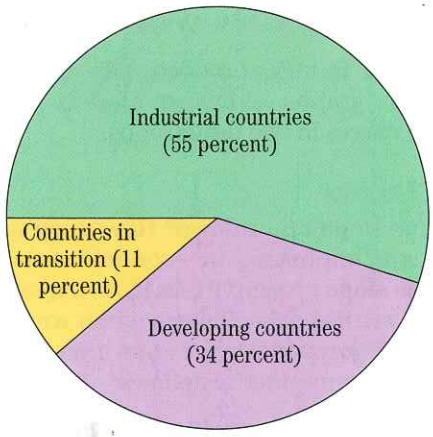


Figure 2A.8
Pie Chart Showing the Shares of the World's GDP
The pie chart shows how the world's GDP is divided up into that produced by (1) the industrial countries, such as the United States, Germany, and Japan; (2) the developing countries, such as India, China, and Nigeria; and (3) countries in transition from communism to capitalism, such as Russia and Poland.

Scatter Plots

Finally, two variables can be usefully compared with a **scatter plot.** The Cartesian coordinate method is used, as in the time-series graph; however, we do not put the year on one of the axes. Instead, the horizontal axis is used for one of the variables and the vertical axis for the other variable. We do this for the debt percentage and the interest rate in Figure 2A.7. The interest rate is on the vertical axis, and the debt percentage is on the horizontal axis. For example, the point at the upper left is 26.8 percent for the debt as a percentage of GDP and 11.5 percent for the interest rate.

Pie Charts

Time-series graphs, bar graphs, and scatter plots are not the only visual ways to observe economic data. For example, the *pie chart* in Figure 2A.8 is useful for comparing percentage shares for a small number of different groups or a small number of time periods. In this example, the pie chart is a visual representation of how the industrial countries produce more than half of the world's GDP, while the developing countries produce 34 percent and the former communist countries in Eastern Europe and the former Soviet Union, now in transition toward market economies, produce about 11 percent.

Visualizing Models with Graphs

Graphs can also represent models. Like graphs showing observations, graphs showing models are usually restricted to curves in two dimensions.

Slopes of Curves

Does a curve slope up or down? How steep is it? These questions are important in economics, as in other sciences. The **slope** of a curve tells us how much the variable on the vertical axis changes when we change the variable on the horizontal axis by one unit.

The slope is computed as follows:

$$\text{Slope} = \frac{\text{change in variable on vertical axis}}{\text{change in variable on horizontal axis}}$$

In most algebra courses, the vertical axis is usually called the y-axis and the horizontal axis is called the x-axis. Thus, the slope is sometime described as

$$\text{Slope} = \frac{\text{change in } y}{\text{change in } x} = \frac{\Delta y}{\Delta x}$$

where the Greek letter Δ (delta) means "change in." In other words, the slope is the ratio of the "rise" (vertical change) to the "run" (horizontal change).

Figure 2A.9 shows how to compute the slope. In this case, the slope declines as the variable on the x-axis increases.

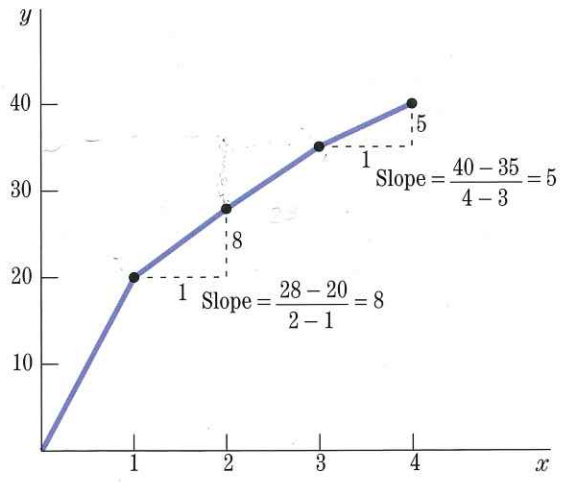


Figure 2A.9
Measuring the Slope
The slope between two points is given by the change along the vertical axis divided by the change along the horizontal axis. In this example, the slope declines as x increases. Since the curve slopes up from left to right, it has a positive slope.

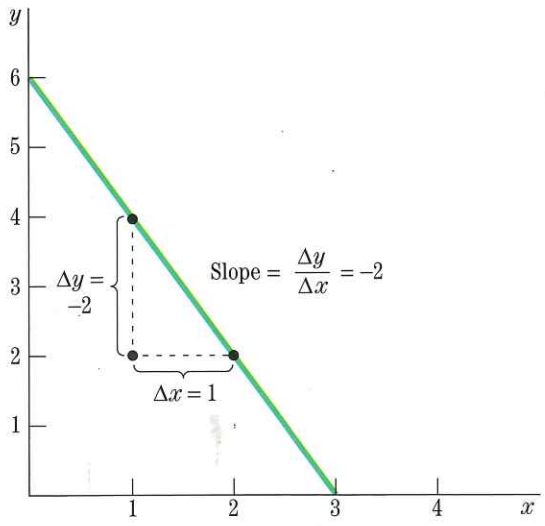


Figure 2A.10
A Relationship with a Negative Slope
Here the slope is negative: $(\Delta y)/(\Delta x) = -2$. As x increases, y falls. The line slopes down from left to right. In this case, y and x are inversely, or negatively, related.

Observe that *the steeper the curve, the larger the slope.* When the curve gets very flat, the slope gets close to zero. Curves can either be upward-sloping or downward-sloping. If the curve slopes up from left to right, as in Figure 2A.9, it has a **positive slope,** and we say that the two variables are positively related. If the curve slopes down from left to right, it has a **negative slope,** and we say that the two variables are negatively related. Figure 2A.10 shows a case where the slope is negative. When x increases by 1 unit ($\Delta x = 1$), y declines by 2 units ($\Delta y = -2$). Thus, the slope equals -2; it is negative. Observe how the curve slopes down from left to right.

If the curve is a straight line, then the slope is a constant. Curves that are straight lines—as in Figure 2A.10—are called **linear.** But economic relationships do not need to be linear, as the example in Figure 2A.9 makes clear. Figure 2A.11 shows six different examples of curves and indicates how they are described.

Graphs of Models with More than Two Variables

In most cases, economic models involve more than two variables. For example, the number of physical examinations could depend on the number of nurses as well as the number of doctors. Or the amount of lemonade demanded might depend on the weather as well as on the price.

Economists have devised several methods for representing models with more than two variables with two-

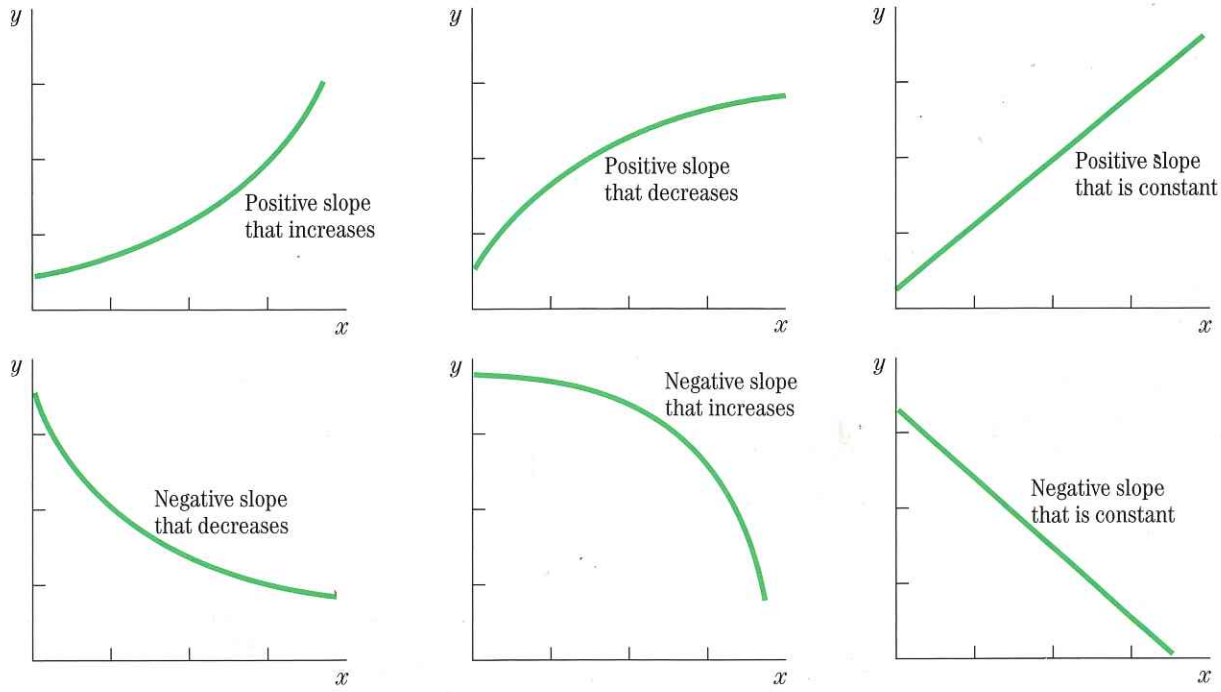


Figure 2A.11
Six Types of Relationships
In the top row, the variables are positively related. In the bottom row, they are negatively related.

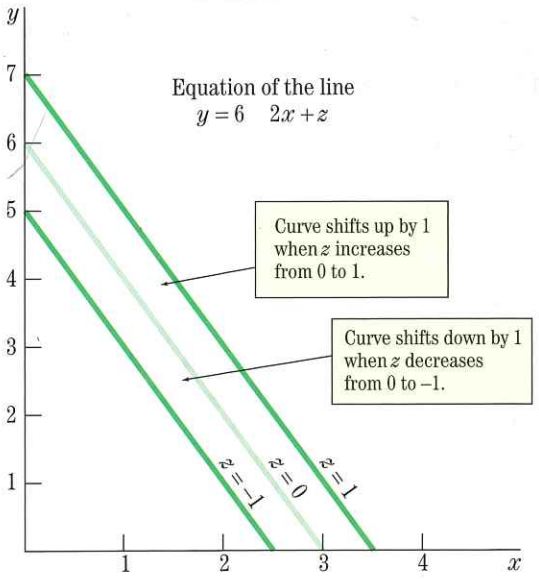


Figure 2A.12
A Third Variable Shifts the Curve
In order to represent models with three variables (x, y, and z) on a two-dimensional graph, economists distinguish between movements along the curve (when x and y change, holding z unchanged) and shifts of the curve (when z changes).

dimensional graphs. Suppose, for example, that the relationship between y and x in Figure 2A.10 depends on a third variable z. For a given value of x, larger values of z lead to larger values of y. This example is graphed in Figure 2A.12. As in Figure 2A.10, when x increases, y falls. This is a **movement along the curve.** But what if z changes? We represent this as a **shift of the curve.** An increase in z shifts the curve up; a decrease in z shifts the curve down.

Thus, by distinguishing between shifts of and movements along a curve, economists represent models with more than two variables in only two dimensions. Only two variables (x and y) are shown explicitly on the graph, and when the third (z) is fixed, changes in x and y are movements along the curve. When z changes, the curve shifts. The distinction between "movements along" and "shifts of" curves comes up many times in economics.

Key Terms and Definitions

Cartesian coordinate system: a graphing system in which ordered pairs of numbers are represented on a plane by the distances from a point to two perpendicular lines, called axes.

time-series graph: a graph that plots a variable over time, usually with time on the horizontal axis.

dual scale: a graph that uses time on the horizontal axis and different scales on the left and right vertical axes to compare the movements of two variables over time.

scatter plot: a graph in which points in a Cartesian coordinate system represent the values of two variables.

slope: a characteristic of a curve that is defined as the change in the variable on the vertical axis divided by the change in the variable on the horizontal axis.

positive slope: a slope of a curve that is greater than zero, representing a positive or direct relationship between two variables.

negative slope: a slope of a curve that is less than zero, representing a negative or inverse relationship between two variables.

linear: a situation in which a curve is straight, with a constant slope.

movement along the curve: a situation in which a change in the variable on one axis causes a change in the variable on the other axis, but the position of the curve is maintained.

shift of the curve: a change in the position of a curve, usually caused by a change in a variable not represented on either axis.

Questions for Review

1. What is the difference between a scatter plot and a time-series graph?
2. Why are dual scales sometimes necessary?
3. What is the advantage of graphs over verbal representations of models?
4. What does a curve with a negative slope look like?
5. What is the difference between a shift in a curve and a movement along a curve?

Problems

1. The table below presents data on the debt and the debt to GDP ratio predicted by the Congressional Budget Office for the United States for each year through 2015.
   a. Construct a time-series plot of the ratio of government debt to GDP.
   b. Construct a time-series plot of the debt.
   c. Construct a scatter plot of the debt ratio and the debt.

Year	Debt	Debt to GDP Ratio
2005	4,656	38.1
2006	4,965	38.5
2007	5,246	38.6
2008	5,506	38.5
2009	5,737	38.2
2010	5,949	37.8
2011	6,054	36.7
2012	6,004	34.8
2013	5,941	33.0
2014	5,847	31.1
2015	5,726	29.1

Source: Congressional Budget Office.

2. The following table presents data on U.S. turkey production and prices.

Year	Turkey Production (billions of pounds)	Price per Pound
1995	5.07	0.99
1996	5.40	1.02
1997	5.41	0.98
1998	5.22	0.95
1999	5.23	0.98
2000	5.33	0.99

   a. Construct a time-series plot of turkey production in the United States.
   b. Construct a time-series plot of the price of turkey per pound.
   c. Construct a scatter plot of turkey production and turkey prices.

3. The following table shows the number of physical examinations given by doctors at an HMO with three different-size clinics: small, medium, and large. The larger the clinic, the more patients the doctors can handle.

Exams per Small Clinic	Exams per Medium Clinic	Exams per Large Clinic	Number of Doctors
0	0	0	0
20	30	35	1
28	42	49	2
35	53	62	3
40	60	70	4

   a. Show the relationship between doctors and physical exams given with *three* curves, where the number of doctors is on the horizontal axis and the number of examinations is on the vertical axis.
   b. Describe how the three relationships compare with one another.
   c. Is a change in the number of doctors a shift of or a movement along the curve?
   d. Is a change in the size of the clinic a shift of or a movement along the curve?

or biologists, need to systematically document and quantify their observations and look for patterns. If we can establish the date when dinosaurs became extinct, then we may be able to test our hunch that a cataclysmic event such as an asteroid hitting the earth caused their extinction. To illustrate how economists document and quantify their observations, let us briefly focus on the last two of the preceding questions, the ones concerning health care and the economy. This will also give us an opportunity to introduce some key indicators used to measure the economy.

Health-Care Spending in America

Health-care spending and prices are more than a curiosity. The more a society spends on health care, the less it can spend on other things, as we know from the production possibilities curve of Chapter 1. Concerns about health-care spending have led to major proposals for changing the way health care is provided and paid for in the United States. Health care *is* a major political issue. People who are dissatisfied with their own health care want the right to sue their health-care providers. Debates about the addition of a prescription drug benefit to Medicare—a government health-care program for the elderly—raged during the period leading up to the 2004 presidential and congressional elections. But let's focus on our first observation and question.

Observation 1: Health-care spending has increased faster than the rest of the U.S. economy since 1990.

How has health-care spending changed relative to the rest of the economy? To determine this, we need a measure of health-care spending and a measure of the size of the overall economy.

Spending as a Share of GDP

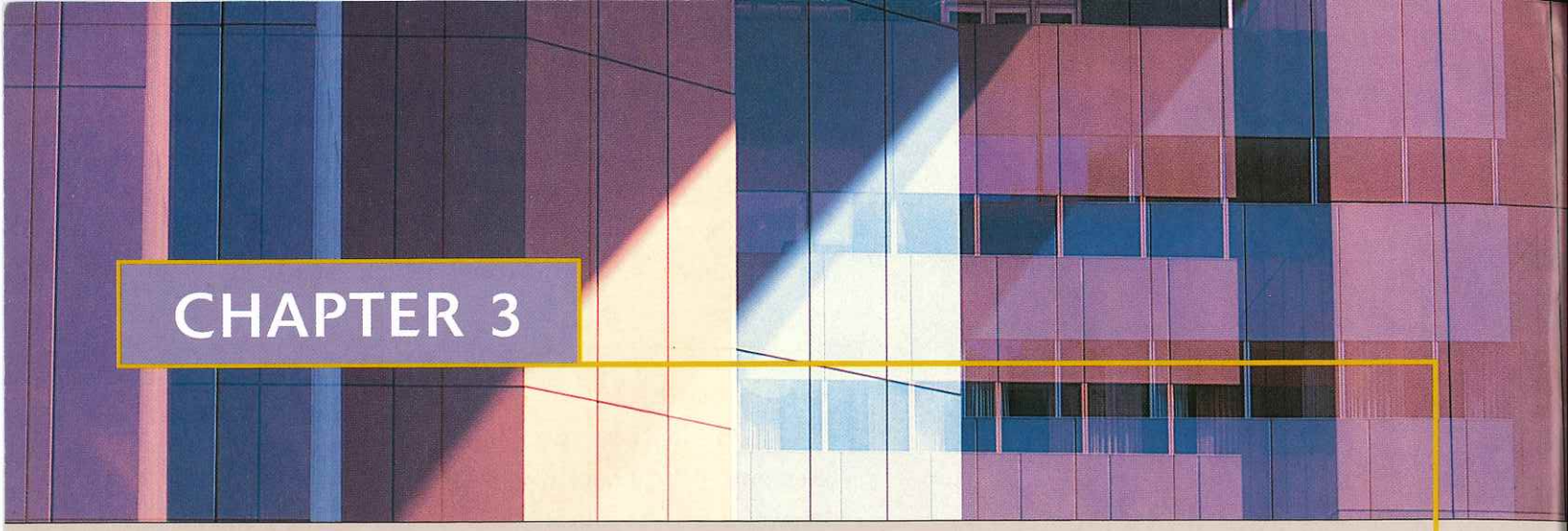
gross domestic product (GDP): a measure of the value of all the goods and services newly produced in an economy during a specified period of time.

The most comprehensive available measure of the size of an economy is the **gross domestic product (GDP).** For the United States, GDP is the total value of all products made in the United States during a specified period of time, such as a year. GDP includes all newly made goods, such as cars, trucks, shoes, airplanes, houses, and telephones; it also includes services, such as education, rock concerts, and health care. To measure the total value of all products made in the economy, economists add up the dollars that people spend on the products.

How large is GDP in the United States? In 2004, it was $11,735 billion, or about $11.7 trillion. We can compute GDP for any year. The question about health-care spending and the size of the economy requires that we look at the U.S. economy since 1990. In 1990, GDP was $5,803 billion. Column (1) of Table 2.1 provides a history of GDP since 1990.

Graphs are frequently a more helpful way to present data like those shown in Table 2.1. Figure 2.1 plots the data on GDP from column (1) of Table 2.1. The vertical axis is measured in billions of dollars; the horizontal axis is measured in years. For example, the point at the extreme lower left in Figure 2.1 represents GDP of $5,803 billion (on the vertical axis) in the year 1990 (on the horizontal axis). The points are connected by a line, which helps us visualize the steady growth of GDP during this period.

Now let us consider health-care spending, which includes payments for hospital services, lab tests, nursing homes, visits to the doctor or dentist, drugs, hearing aids, and eyeglasses. If we add up all spending on health care, we get $1,392 billion in 2004.



CHAPTER 3

The Supply and Demand Model

It's pretty much the same thing every March. Four college basketball teams win a place in the top round of the national tournament, the Final Four. Each college lets students at the college buy a limited number of tickets for about $100 a seat. Then, when the students get to the city where the Final Four is being played, they find people on the street willing to pay staggering amounts of money for those tickets—as much as $5,000. And it is always tempting to sell the tickets and watch the game on TV. Some, of course, do sell. But how are these prices on the street market (and, increasingly, through ticket brokers on the Internet) determined? How are prices in general determined? What causes the price of health care to rise? What causes the price of computers to fall? What determines the price at which people buy or sell gasoline, electronic goods, printing services, or foreign currencies? The purpose of this chapter is to show how to find the answers to such questions.

To do so, we need to construct a model—a simplified description of how a market works. The model economists use to explain how prices are determined in a market is called the *supply and demand model.* This model describes how particular markets—such as the health-care market or the computer market—work. It consists of three elements: *demand,* describing the behavior of consumers in the market; *supply,* describing the behavior of firms in the market; and *market equilibrium,* connecting supply and demand and describing how consumers and firms interact in the market.

Supply and Demand in the Final Four
The model of supply and demand can explain how $500 tickets to the NCAA Final Four are sold well before the first college basketball game of the season. It also explains more routine buyer and seller interactions, describing the behavior of sellers and buyers and how they connect in markets.

Economists like to compare the supply and demand model to a pair of scissors. Demand is one blade of the scissors. Supply is the other. Either blade alone is incomplete and virtually useless; but when the two blades of a pair of scissors are connected to form the scissors, they become an amazingly useful, yet simple, tool. So it is with the supply and demand model.

In this chapter, we first describe each of the three elements of the model. We then show how to use the model to answer a host of questions about price determination in a market economy. The news story excerpted in the "Reading the News About . . ." box on page 52 illustrates how supply and demand works in a real (world oil) market. Think about the issues considered there as you learn the elements of the model. The Case Study near the end of this chapter takes a closer look at how supply and demand works in a related real world market—the gasoline market.

Demand

demand: a relationship between **price** and **quantity demanded.**

price: the amount of money or other goods that one must pay to obtain a particular good.

quantity demanded: the quantity of a good that people want to buy at a given price during a specific time period.

To an economist, the term *demand*—whether the demand for health care or the demand for computers—has a very specific meaning. **Demand** is a relationship between two economic variables: (1) *the price of a particular good* and (2) *the quantity of the good consumers are willing to buy at that price during a specific time period,* all other things being equal. For short, we call the first variable the **price** and the second variable the **quantity demanded.** The phrase *all other things being equal,* or *ceteris paribus,* is appended to the definition of demand because the quantity consumers are willing to buy depends on many other things besides the price of the good; we want to hold these other things constant, or equal, while we examine the relationship between price and quantity demanded.

Demand can be represented with a numerical table or a graph. In either case, demand describes how much of a good consumers will purchase at each price. Consider the demand for bicycles in the United States. An example of the demand for bicycles is shown in Table 3.1. Several prices for a typical bicycle are listed in the first

Table 3.1
Demand Schedule for Bicycles (millions of bicycles per year)

Price	Quantity Demanded	Price	Quantity Demanded
$140	18	$240	5
$160	14	$260	3
$180	11	$289	2
$200	9	$300	1
$220	7		

Economists use the supply and demand model both to explain past observations and to better understand what would happen to prices in a particular market under different scenarios. The hypothetical example of a bicycle market has been useful for explaining general features of the supply and demand model. But now we want to show how the model can be applied to real-world situations. To illustrate the applicability of the supply-demand model, we look in detail at a specific market—the market for gasoline in the United States. Those of you who own cars are keenly aware of the impact of recent price fluctuations in this market.

Bush Touts Energy Technologies To Satisfy Global, U.S. Demand

By John D. McKinnon and
John J. Fialka
Staff Reporters
WALL STREET JOURNAL
APRIL 28, 2005

President Bush sought to dispel growing worries about energy costs, saying new technologies promise eventually to end U.S. dependence on pricey foreign oil and help meet soaring global energy demands in places like China and India.

In an upbeat speech to a generally friendly audience of small-business owners, Mr. Bush outlined a series of measures he plans to expand U.S. energy capacity, from promoting new nuclear power-plant construction to building new oil refineries on old military bases. He also wants to encourage use of new and alternative fuels, including cleaner-burning diesel soon to come on the market, and promote a range of new conservation technologies.

Mr. Bush also said his administration will push for more use of clean-coal and nuclear technologies to answer growing energy needs in fast-growing economies in Asia. That rising demand is contributing substantially to current high prices for oil and gasoline, but Mr. Bush said the trend is on the way to being reversed.

"Our country is on the doorstep of incredible technological advances that will make energy more abundant and more affordable for our citizens," the president said. "By harnessing the power of technology, we're going to be able to grow our economy, protect our environment, and achieve greater energy independence. That's why I'm so optimistic about our future here in America."

Mr. Bush and his aides have been at pains for several days to show that current high gasoline prices are only part of an energy picture that is actually brighter than critics say—and will improve further if Congress finally passes Mr. Bush's long-delayed energy proposals.

Earlier this week, Mr. Bush invited Saudi Crown Prince Abdullah to his ranch in Texas to highlight that country's plans to expand its oil-producing capacity. Erosion of the Saudis' capacity cushion has been a big reason for the recent spike in the crude oil prices.

Yesterday's speech was aimed at filling in the rest of the picture.

Democratic leaders were predictably skeptical it would have much short-term impact. Sen. John Kerry (D., Mass.) said the Republicans' energy plans "won't help the families, truckers, farmers and small businesses suffering from skyrocketing gas prices today."

David Hamilton, a spokesman for the Sierra Club, asserted that the president's speech was a continuation of administration policies that "help industry, step on state's rights and thumb a nose at sound science."

Business groups were more enthusiastic. Allen Schaeffer, executive director of the Diesel Technology Forum, says Mr. Bush's plans for a new tax credit to encourage consumers to buy clean-diesel cars will help auto makers reach the decision that the U.S. market has more potential than previously thought. Mr. Schaeffer noted that tax incentives in Western Europe have convinced 42% of new-car buyers to buy diesels. By contrast, only 3.3% of new cars and light trucks in the U.S. are diesel powered, and most of them are pickup trucks.

John Rice, president and CEO of GE Energy, which is building clean-coal and nuclear plants, said the new initiatives represent a big opportunity for his company, both domestically and abroad. "The president understands clearly that all of us have to be in the game in these new technologies in order for them to be effective and have an impact," he said.

The need to reduce demand

The need to increase supply

Source: Republished with permission of the *Wall Street Journal*, from John C. McKinnon and John J. Fialka, "Bush Touts Energy Technologies to Satisfy Global, US Demand," April 28, 2005, conveyed via Copyright Clearance Center, Inc.

column of the table, ranging from $140 to $300. Of course, there are many kinds of bicycles—mountain bikes, racing bikes, children's bikes, and inexpensive one-speed bikes with cruiser brakes—so you need to think about the price of an average, or typical, bike.

Listed in the second column of Table 3.1 is the quantity demanded (in millions of bicycles) each year in the United States at the price in the first column. This is the total demand in the bicycle market. For example, at a price of $180 per bicycle, consumers would buy 11 million bicycles. That is, the quantity demanded would be 11 million bicycles each year in the United States, according to Table 3.1.

Price	Quantity Demanded
$140	18
$160	14
$180	11
$200	9
$220	7
$240	5
$260	3
$280	2
$300	1

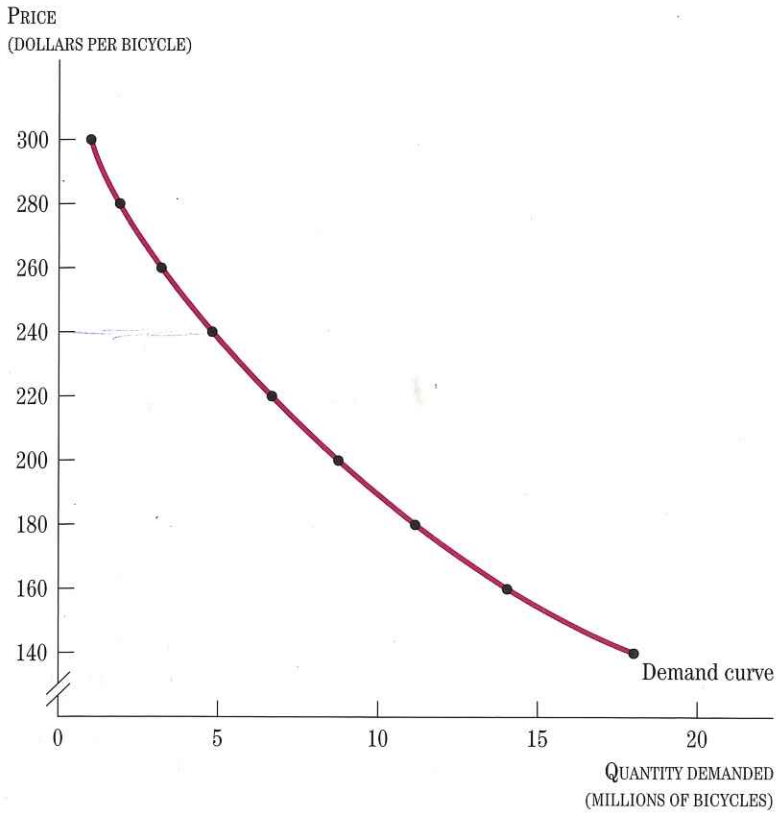


Figure 3.1
The Demand Curve

The demand curve shows that the price of a good and the quantity demanded by consumers are negatively related. The curve slopes down. For each price, the demand curve gives the quantity demanded, or the quantity that consumers are willing to buy at that price. The points along the demand curve for bicycles shown here are the same as the pairs of numbers in Table 3.1.

Observe that as the price rises, the quantity demanded by consumers goes down. If the price goes up from $180 to $200 per bicycle, for example, the quantity demanded goes down from 11 million to 9 million bicycles. On the other hand, if the price goes down, the quantity demanded goes up. If the price falls from $180 to $160, for example, the quantity demanded rises from 11 million to 14 million bicycles.

The relationship between price and quantity demanded in Table 3.1 is called a **demand schedule.** This relationship is an example of the law of demand. The **law of demand** says that the higher the price, the lower the quantity demanded in the market; and the lower the price, the higher the quantity demanded in the market. In other words, the law of demand says that the price and the quantity demanded are negatively related, all other things being equal.

demand schedule: a tabular presentation of demand showing the price and quantity demanded for a particular good, all else being equal.

law of demand: the tendency for the quantity demanded of a good in a market to decline as its price rises.

The Demand Curve

Figure 3.1 represents demand graphically. It is a graph with the price of the good on the vertical axis and the quantity demanded of the good on the horizontal axis. It shows the demand for bicycles given in Table 3.1. Each of the nine rows in Table 3.1

corresponds to one of the nine points in Figure 3.1. For example, the point at the lower right part of the graph corresponds to the first row of the table, when the price is $140 and the quantity demanded is 18 million bicycles. The resulting curve showing all the combinations of price and quantity demanded is the **demand curve.** It slopes downward from left to right because the quantity demanded is negatively related to the price. To remember that the *d*emand curve slopes *d*ownward, think of the *d* in *demand*.

demand curve: a graph of demand showing the downward-sloping relationship between price and quantity demanded.

Why does the demand curve slope downward? The demand curve tells us the quantity demanded by all consumers. Consumers must make choices with scarce resources. They must choose between bicycles and other goods. If the price of bicycles falls, then some consumers who previously found the price of bicycles too high may decide to buy a bicycle. The lower price of bicycles gives them an incentive to buy bicycles rather than other goods. It is important to remember that when economists draw a demand curve, they hold constant the price of other goods: running shoes, in-line skates, motor scooters, and so on. When the price of bicycles falls, bicycles become more attractive to people in comparison with these other goods. As a result, the quantity demanded rises when the price falls. Conversely, when the price of bicycles rises, some people may decide to buy in-line skates or motor scooters instead of bicycles. As a result, the quantity demanded declines when the price rises.

There's plenty of evidence in the real world that demand curves are downward-sloping. In June of 2004, vehicle sales at General Motors were slowing. In July, General Motors increased the cash-back offer on most of its trucks and cars. You might (correctly) speculate that this reduction in the price of vehicles was intended to increase vehicle sales. The policy implications of a downward-sloping demand curve frequently sound like common sense. Suggestions on how to reduce student drinking on college campuses include raising the price of alcohol. The idea, of course, is that students would buy less alcohol if it were more expensive.

Shifts in Demand

Now, price is not the only thing that affects the quantity of a good that people buy. The weather, people's concerns about the environment, or the availability of bike lanes on roads can influence people's decisions to purchase bicycles, for example. The quantity of bicycles bought might increase if a climate change brought on an extended period of dry weather. Because people would enjoy riding their bicycles more in dry weather, more bicycles would be purchased at any given price. Or perhaps a health trend might lead people to get exercise by riding bicycles rather than driving their cars. This would also lead to more purchases of bicycles.

The demand curve is drawn assuming that all other things are equal, except the price of the good. A change in any one of these other things, previously assumed to be equal, will shift the demand curve. An increase in demand shifts the demand curve to the right. A decrease in demand shifts the demand curve to the left. This is illustrated in Figure 3.2. The lightly shaded curve labeled "old demand curve" is the same as the demand curve in Figure 3.1. The arrow shows how this curve has shifted to the right to the more darkly shaded curve labeled "new demand curve." Thus, Figure 3.2 shows the demand curve for bicycles shifting to the right. When the demand curve shifts to the right, more bicycles are purchased than before at any given price. For example, before the shift in demand, a $200 price led to 9 million bicycles purchased. But when the demand curve shifts to the right because of drier weather, that same price leads to 13 million bicycles purchased. The demand curve would shift to the left if a climate change to wetter weather reduced people's purchases of bicycles at any given price.

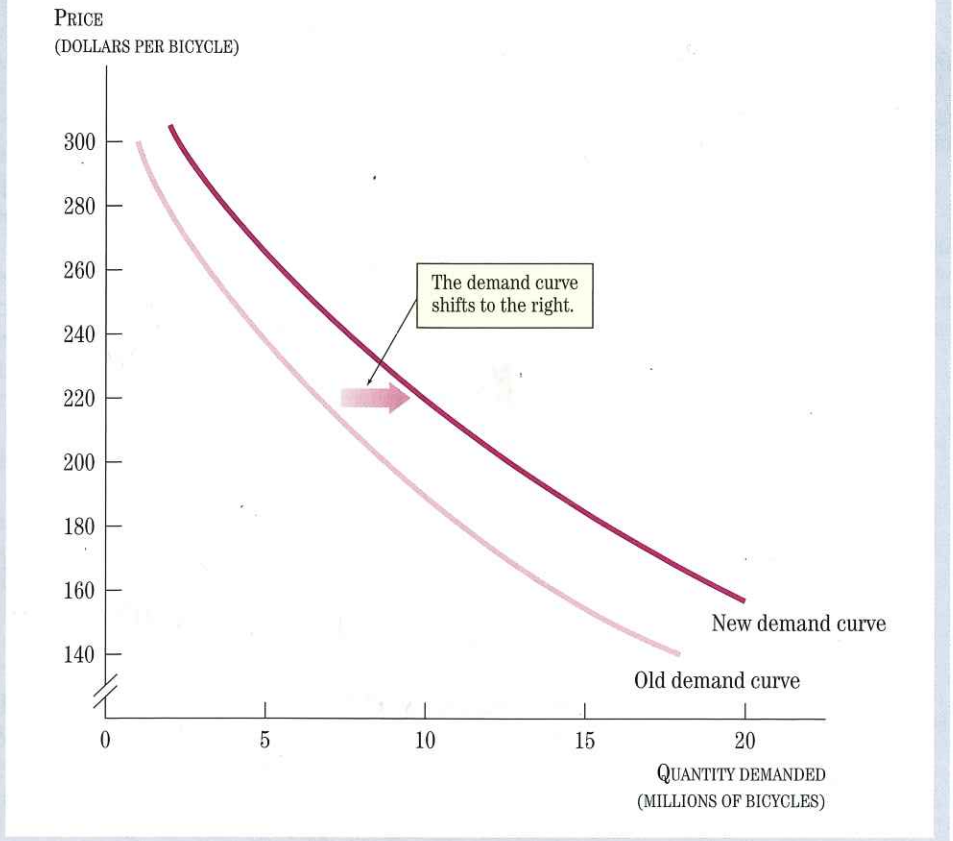


Figure 3.2
A Shift in the Demand Curve
The demand curve shows how the quantity demanded of a good is related to the price of the good, all other things being equal. A change in one of these other things—the weather or people's tastes, for example—will shift the demand curve, as shown in the graph. In this case, the demand for bicycles increases; the demand curve for bicycles shifts to the right.

There are many reasons the demand curve may shift. Most of them can be attributed to one of several sources: *consumers' preferences, consumers' information, consumers' incomes, the number of consumers in the market, consumers' expectations of future prices,* and *the price of related goods.* Let us briefly consider each source of shifts in demand.

■ **Consumers' Preferences.** In general, a change in people's tastes or preferences for a product compared to other products will change the amount of the product they purchase at any given price. After September 11, 2001, changes in consumers' preferences shifted the demand for many goods. Examples include an increase in the demand for stuffed animals as adults searched for comfort, an increase in New York City in the demand for shoes that are easy to walk in, and an increase in the demand for cell phones by people who wanted more access to communication.

■ **Consumers' Information.** A change in information relating to a product can also cause the demand curve to shift. For example, when people learned about the dangers of smoking, the demand for cigarettes declined. Immediately after the mass recall of Firestone tires in 2000, there was a decrease in demand for Firestone tires, which became evident when the sales of Firestone tires fell by 40 percent.

normal good: a good for which demand increases when income rises and decreases when income falls.

inferior good: a good for which demand decreases when income rises and increases when income falls.

substitute: a good that has many of the same characteristics as and can be used in place of another good.

■ **Consumers' Incomes.** If people's incomes change, then their purchases of goods usually change. An increase in income increases the demand for most goods. A decline in income reduces the demand for these goods. Goods for which demand increases when income rises and decreases when income falls are called **normal goods** by economists. For example, when the U.S. stock market boomed in the late 1990s, demand for jewelry increased as a result of the increase in consumers' wealth. The economic slowdown in 2001 caused a decrease in demand for snowplow service. Jewelry and snowplow service are therefore examples of normal goods.

However, the demand for some goods may decline when income increases. Such goods are called **inferior goods** by economists. The demand for inferior goods declines when people's income increases because they can afford more attractive goods. During the late 1990s U.S. stock market boom, demand for Christmas candy decreased. The economic slowdown in 2001 caused an increase in demand for snow shovels. This tells you that Christmas candy and snow shovels are inferior goods. Jewelry is more attractive to consumers than candy, and snowplow service is more attractive to consumers than snow shovels.

Notice that whether goods are normal or inferior, the demand for them usually shifts when consumers' incomes change. In 2001, as a result of the economic slowdown, lipstick sales increased. Does this tell you that lipstick is a normal or an inferior good? Since demand increased as a result of a decrease in consumers' incomes, lipstick is an inferior good. One explanation is that women were suddenly more careful with their money and bought lipstick because of its low price, even though other more expensive items were more attractive.

■ **Number of Consumers in the Market.** Demand is a relationship between price and the quantity demanded by *all* consumers in the market. If the number of consumers increases, then demand will increase. If the number of consumers falls, then demand will decrease. For example, the number of teenagers in the U.S. population expanded sharply in the late 1990s. This increased the demand for *Seventeen* magazine, for Rollerblades, for Clearasil, and for other goods that teenagers tend to buy. As the baby-boom generation in the United States ages, the demand for home hair coloring kits and luxury skin care products (for men as well as for women) is increasing.

■ **Consumers' Expectations of the Future Price.** If people expect the price of a good to increase, they will want to buy it before the price increases. Conversely, if people expect the price to decline, they will purchase less and wait for the decline. One sees this effect of expectations of future price changes often. "We'd better buy before the price goes up" is a common reason for purchasing items during a clearance sale. Or, "Let's put off buying that big-screen TV until the postholiday sales."

In general, it is difficult to forecast the future, but sometimes consumers know quite a bit about whether the price of a good will rise or fall, and they react accordingly. Thus, demand increases if people expect the *future* price of the good to rise. And demand decreases if people expect the *future* price of the good to fall.

In 1995, President Clinton threatened a 100 percent tariff (tax) on some luxury cars produced in Japan. This resulted in an immediate increase in demand for these cars, since buyers were afraid they would become too expensive after the tariff was imposed.

■ **Prices of Closely Related Goods.** A change in the price of a closely related good can increase or decrease demand for another good, depending on whether the good is a substitute or a complement. A **substitute** is a good that provides some of

Substitutes and complements *Music CDs and downloaded music are examples of substitutes; they share similar characteristics. You would expect, therefore, that a rise in the price of CDs would result in an increase in the sale of downloaded music- and vice versa. SUVs and gasoline are examples of complements; they tend to be consumed together. With an increase in gasoline prices in 2004 and 2005, consumers were less eager to purchase SUVs, and their sales declined.*

complement: a good that is usually consumed or used together with another good.

the same uses or enjoyment as another good. Butter and margarine are substitutes. In general, the demand for a good will increase if the price of a substitute for the good rises, and the demand for a good will decrease if the price of a substitute falls. Sales of CDs and downloaded music are substitutes. You would therefore expect an increase in the price of downloaded music to increase the demand for CDs. This may help explain why the recording industry filed lawsuits against users of online file-sharing software in 2003.

A **complement** is a good that tends to be consumed together with another good. Gasoline and SUVs are complements. The increase in gasoline prices in 2004 and 2005 led to a decrease in demand for SUVs.

Movements Along versus Shifts of the Demand Curve

We have shown that the demand curve can shift, and we have given many possible reasons for such shifts. In using demand curves, it is very important to distinguish *shifts* of the demand curve from *movements along* the demand curve. This distinction is illustrated in Figure 3.3.

A *movement along* the demand curve occurs when the quantity demanded changes as a result of a *change in the price of the good.* For example, if the price of bicycles rises, causing the quantity demanded by consumers to fall, then there is a movement along the demand curve. A movement along the demand curve for bicycles occurs when the quantity demanded changes from point *A* to point *B* or from point *A* to point *C* in Figure 3.3. At point *A*, the price is $200 and the quantity demanded is 9 million. At point *B*, the price is $220 and the quantity demanded is 7 million. If the quantity changes because the price changes, economists say that there is a *change in the quantity demanded.*

A *shift* of the demand curve, on the other hand, occurs if there is a change due to *any source except the price.* When the demand curve shifts, economists say that there is a *change in demand.* Remember, the term *demand* refers to the entire curve or schedule relating price and quantity demanded, while the term *quantity demanded* refers to a single point on the demand curve. You should be able to tell whether any economic event causes (1) a change in demand or (2) a change in the quantity demanded; or, equivalently, (1) a shift in the demand curve or (2) a movement along the demand curve.

Here's an example to test your understanding of demand shifts and movement along the demand curve. Disney's theme park attendance was lower in 2001 than in

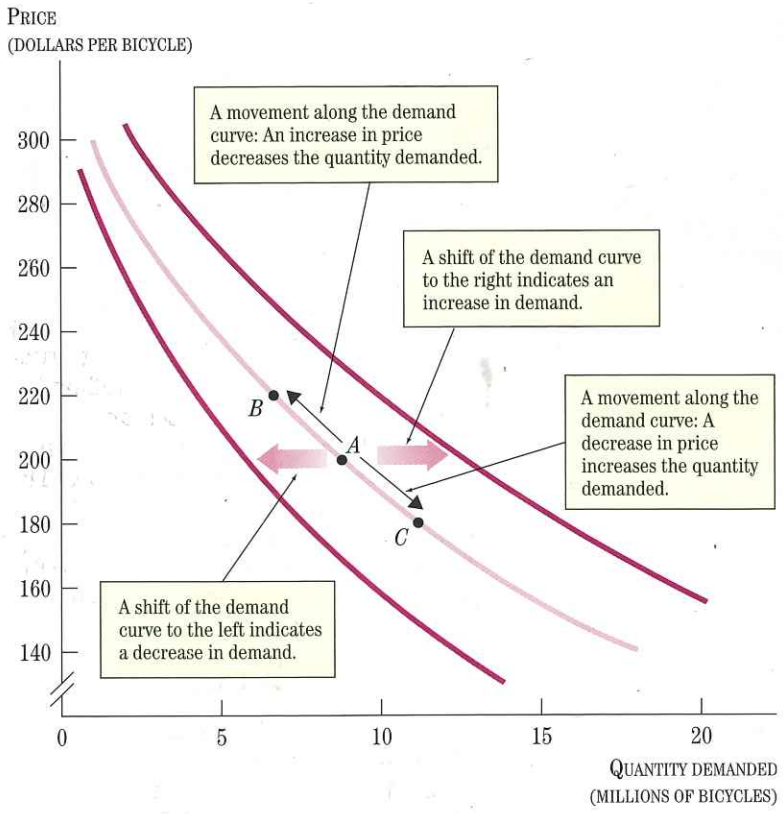


Figure 3.3
Shifts of versus Movements Along the Demand Curve
A *shift* of the demand curve occurs when there is a change in something (other than the good's own price) that affects the quantity of a good that consumers are willing to buy. An increase in demand is a shift to the right of the demand curve. A decrease in demand is a shift to the left of the demand curve. A *movement along* the demand curve occurs when the price of the good changes, causing the quantity demanded to change, as, for example, from point *A* to point *B* or *C*.

previous years as a result of the weak economy. Because of the fall in attendance, Disney lowered the adult admission price at its California Adventure park. Is this describing a *change in demand* or a *change in the quantity demanded* in the market for theme parks? The decrease in attendance caused by the weak economy in 2001 describes a decrease in demand. Going to a theme park is therefore a normal good. When Disney lowered its admission price, it was in anticipation of an increase in quantity demanded—the park management anticipated more attendance at a lower price. The decrease in price and resulting increase in quantity demanded describe movement along the demand curve.

REVIEW

- Demand is a relationship between the price of a good and the quantity people will buy at each price, all other things being equal. The demand curve slopes down. The price and the quantity demanded are negatively related.
- When the price of a good changes, the quantity demanded changes and we have a movement along the demand curve.
- When something other than the price changes and affects demand, there is a shift in the demand curve, or, simply, a change in demand.

Supply

supply: a relationship between **price** and **quantity supplied.**

quantity supplied: the quantity of a good that firms are willing to sell at a given price.

Table 3.2
Supply Schedule for Bicycles (millions of bicycles per year)

Price	Quantity Supplied
$140	1
$160	4
$180	7
$200	9
$220	11
$240	13
$260	15
$280	16
$300	17

supply schedule: a tabular presentation of supply showing the price and quantity supplied of a particular good, all else being equal.

law of supply: the tendency for the quantity supplied of a good in a market to increase as its price rises.

supply curve: a graph of supply showing the upward-sloping relationship between price and quantity supplied.

Whereas demand refers to the behavior of consumers, supply refers to the behavior of firms. The term *supply*—whether it is the supply of health care or the supply of computers—has a very specific meaning for economists. **Supply** is a relationship between two variables: (1) *the price of a particular good* and (2) *the quantity of the good firms are willing to sell at that price*, all other things being the same. For short, we call the first variable the **price** and the second variable the **quantity supplied.**

Supply can be represented with a numerical table or a graph. An example of the supply of bicycles is shown in Table 3.2. Listed in the first column of Table 3.2 is the price of bicycles; the range of prices is the same as for the demand schedule in Table 3.1. The second column lists the quantity supplied (in millions of bicycles) in the entire market by bicycle-producing firms at each price. For example, at a price of $180, the quantity supplied is 7 million bicycles. Observe that as the price increases, the quantity supplied increases, and that as the price decreases, the quantity supplied decreases. For example, if the price rises from $180 to $200, the quantity supplied increases from 7 to 9 million bicycles. The relationship between price and quantity supplied in Table 3.2 is a **supply schedule.** This relationship is an example of the law of supply. The **law of supply** says that the higher the price, the higher the quantity supplied, and the lower the price, the lower the quantity supplied. In other words, the law of supply says that the price and the quantity supplied are positively related.

The Supply Curve

We can represent the supply schedule in Table 3.2 graphically by plotting the price and quantity supplied on a graph, as shown in Figure 3.4. The scales of each axis in Figure 3.4 are exactly the same as those in Figure 3.1, except that Figure 3.4 shows the quantity supplied, whereas Figure 3.1 shows the quantity demanded. Each pair of numbers in Table 3.2 is plotted as a point in Figure 3.4. The resulting curve showing all the combinations of prices and quantities supplied is the **supply curve.** Note that the curve slopes upward: $280 represents a high price, and there the quantity supplied is high—16 million bicycles. If the price is down at $160 a bicycle, then firms are willing to sell only 4 million bicycles.

Why does the supply curve slope upward? Imagine yourself running a firm that produces and sells bicycles. If the price of the bicycles goes up, from $180 to $280, then you can earn $100 more for each bicycle you produce and sell. Given your production costs, if you earn more from each bicycle, you will have a greater incentive to produce and sell more bicycles. If producing more bicycles increases the costs of producing each bicycle, perhaps because you must pay the bike assembly workers a higher wage for working overtime, the higher price will give you the incentive to incur these costs. Other bicycle firms will be thinking the same way. Thus, firms are willing to sell more bicycles as the price rises. Conversely, the incentive for firms to sell bicycles will decline as the price falls. Basically, that is why there is a positive relationship between price and quantity supplied.

When formulating economic policy, it is important to remember this supply relationship. When the price of a good increases, it leads to an increase in the quantity supplied. U.S. agricultural policy guarantees farmers a specific price on certain crops. The U.S. government pays these farmers if the price the crops sell for is too low. In response to this guaranteed higher price, U.S. farmers grow more crops than U.S. consumers buy. This excess production of crops decreases crop prices in international markets. Therefore, U.S. agricultural policy continues to be discussed in international trade talks.

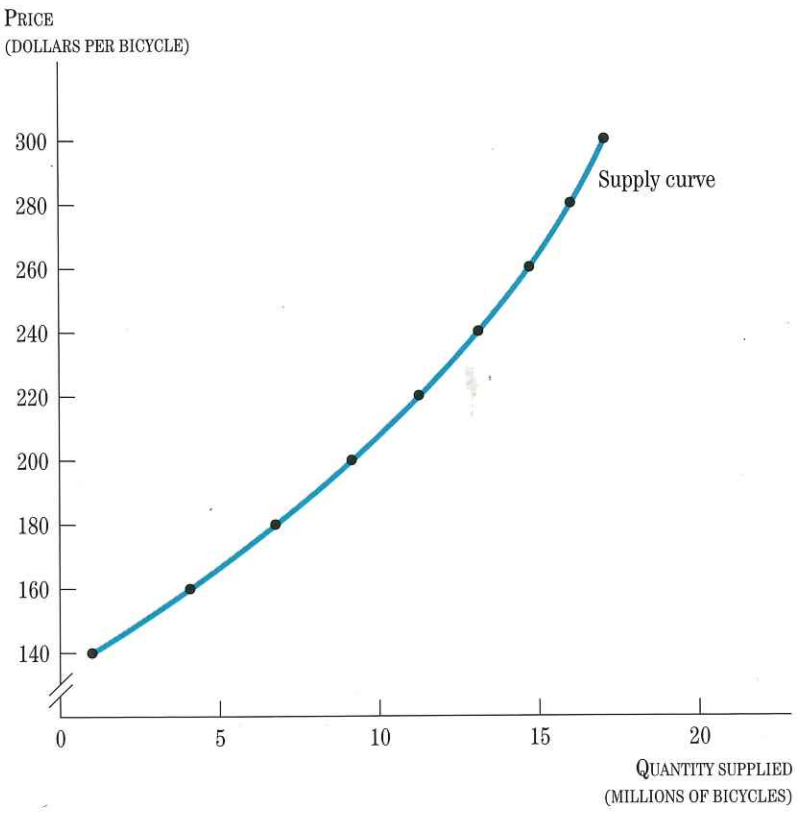


Figure 3.4
The Supply Curve
The supply curve shows that the price and the quantity supplied by firms in the market are positively related. The curve slopes up. For each price on the vertical axis, the supply curve shows the quantity that firms are willing to sell along the horizontal axis. The points along the supply curve for bicycles match the pairs of numbers in Table 3.2.

Shifts in Supply

The supply curve is drawn on the assumption that all other things are equal, except the price of the good. If any one of these other things changes, then the supply curve shifts. For example, suppose a new machine is invented that makes it less costly for firms to produce bicycles; then firms would have more incentive at any given price to produce and sell more bicycles. Supply would increase; the supply curve would shift to the right.

Figure 3.5 shows how the supply curve for bicycles would shift to the right because of a new cost-reducing machine. The supply curve would shift to the left if there were a decrease in supply. Supply would decrease, for example, if bicycle-producing firms suddenly found that their existing machines became too hot and had to be oiled with an expensive lubricant each time a bicycle was produced. This would raise costs, lower supply, and shift the supply curve to the left.

Many things can cause the supply curve to shift. Most of these can be categorized by the source of the change in supply: *technology, the price of goods used in production, the number of firms in the market, expectations of future prices,* and *government taxes, subsidies, and regulations.* Let us briefly consider the sources of shifts in supply.

■ **Technology.** Anything that changes the amount a firm can produce with a given amount of inputs to production can be considered a change in technology. A

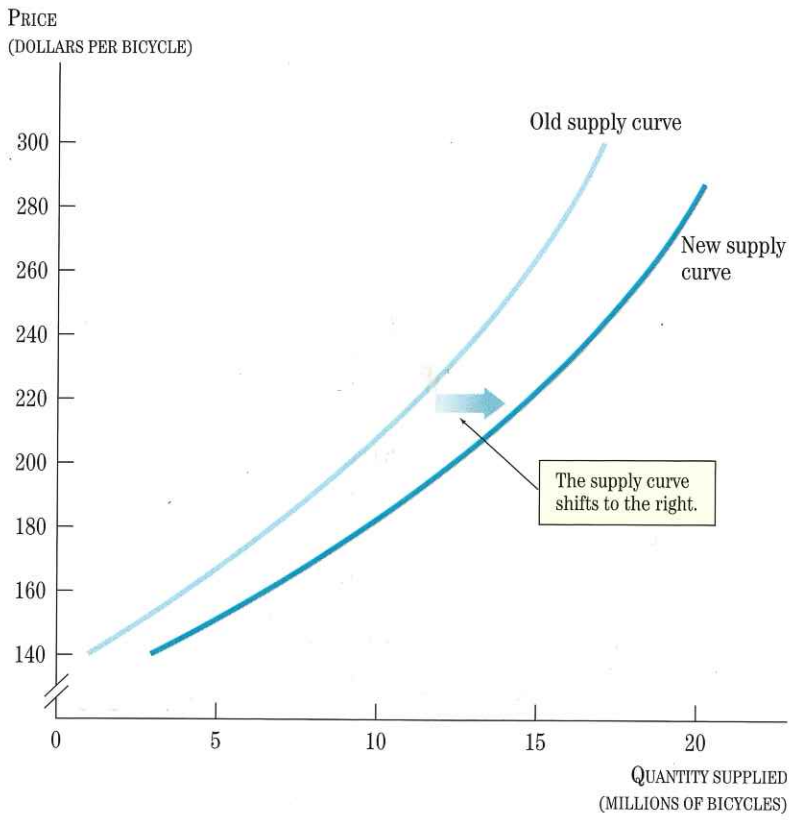


Figure 3.5
A Shift in the Supply Curve
The supply curve is a relationship between the quantity supplied of a good and the price of the good, all other things being equal. A change in one of these other things (other than the good's price) will shift the supply curve, as shown in the graph. In this case, the supply of bicycles increases; the supply curve for bicycles shifts to the right.

2002 study reported the number of labor hours needed to produce an automobile. DaimlerChrysler needed 44 hours per vehicle, while Nissan needed only 29 hours per vehicle. Suppose auto producers initially all needed 44 hours to produce a vehicle and changed their production technology to need instead only 29 hours per vehicle. This improvement in technology would correspond to an increase in supply, a shift in the supply curve to the right. Another way of viewing an increase in supply is that producers are willing to sell any given quantity at a lower price than before. This makes sense, since production costs are lower with the improvement in technology.

Droughts, earthquakes, and terrorists' bombing of factories also affect the amount that can be produced with given inputs. A drought that reduces the amount of wheat that can be produced on a farm in the Midwest and a freeze that reduces the number of oranges yielded by trees in Florida orchards are examples. Because such events change the amount that can be produced with a given amount of inputs, they are similar to changes in technology, though these examples reduce rather than increase supply. In these cases, the supply curve shifts to the left.

■ **The Price of Goods Used in Production.** If the prices of the inputs to production—raw materials, labor, and capital—increase, then it becomes more costly to produce goods, and firms will produce less at any given price; the supply curve will shift to the left. When McDonald's requested in 2000 that its egg suppliers treat their hens more humanely, the egg suppliers were concerned that this would

cause the supply of eggs to decrease. That is, to produce any given quantity of eggs, the price they sold for would have to be higher. An increase in production costs causes supply to decrease, and a decrease in production costs causes supply to increase.

■ **The Number of Firms in the Market.** Remember that the supply curve refers to *all* the firms producing the product. If the number of firms increases, then more goods will be produced at each price; supply increases, and the supply curve shifts to the right.

The same rule applies to the supply of labor. In 2001, a shortage of substitute teachers in many school districts around the country grew to crisis proportions. In response, some school districts decided to lower the requirements for substitute teachers. This, they hoped, would encourage more people to work as substitute teachers at any wage, causing the supply of substitute teachers to increase.

A decline in the number of firms would shift the supply curve to the left. For example, the number of drive-in movie theaters has declined sharply over the last 30 years; hence the supply curve has shifted to the left.

■ **Expectations of Future Prices.** If firms expect the price of the good they produce to rise in the future, then they will hold off selling at least part of their production until the price rises. For example, farmers in the United States who anticipate an increase in wheat prices because of political turbulence in Russia may decide to store more wheat in silos and sell it later, after the price rises. Thus, expectations of *future* price increases tend to reduce supply. Conversely, expectations of *future* price decreases tend to increase supply.

■ **Government Taxes, Subsidies, and Regulations.** The government has the ability to affect the supply of particular goods produced by firms. For example, the government imposes taxes on firms to pay for such government services as education, police, and national defense. These taxes increase firms' costs and reduce supply. The supply curve shifts to the left when a tax on what firms sell in the market increases.

The government also makes payments—subsidies—to firms to encourage the firms to produce certain goods. Such subsidies have the opposite effect of taxes on supply. An increase in subsidies reduces firms' costs and increases the supply. For example, the state of Arizona for a brief time subsidized SUVs that run on both propane and gasoline. Compare this to the U. S. government's imposition of a tax on cigarettes and on alcohol. We interpret the cigarette tax as a decrease in the supply of cigarettes and the hybrid SUV subsidy as an increase in the supply of hybrid SUVs.

Governments also regulate firms. In some cases, such regulations can change the firms' costs of production and thereby affect supply. For example, when the government requires that firms install safety features on their products, the cost of producing the products rises, and thus supply declines.

Movements Along versus Shifts of the Supply Curve

Figure 3.6 compares *shifts* of the supply curve with *movements along* the supply curve. A *movement along* the supply curve occurs when a change in price causes a change in the quantity supplied. Economists then say that there is a *change in the quantity supplied*, as, for example, when the quantity supplied changes from point D to point F or from point D to point E in Figure 3.6.

A *shift* of the supply curve occurs if there is a change due to *any source except the price*. When the supply curve shifts, economists say that there is a *change in supply*.

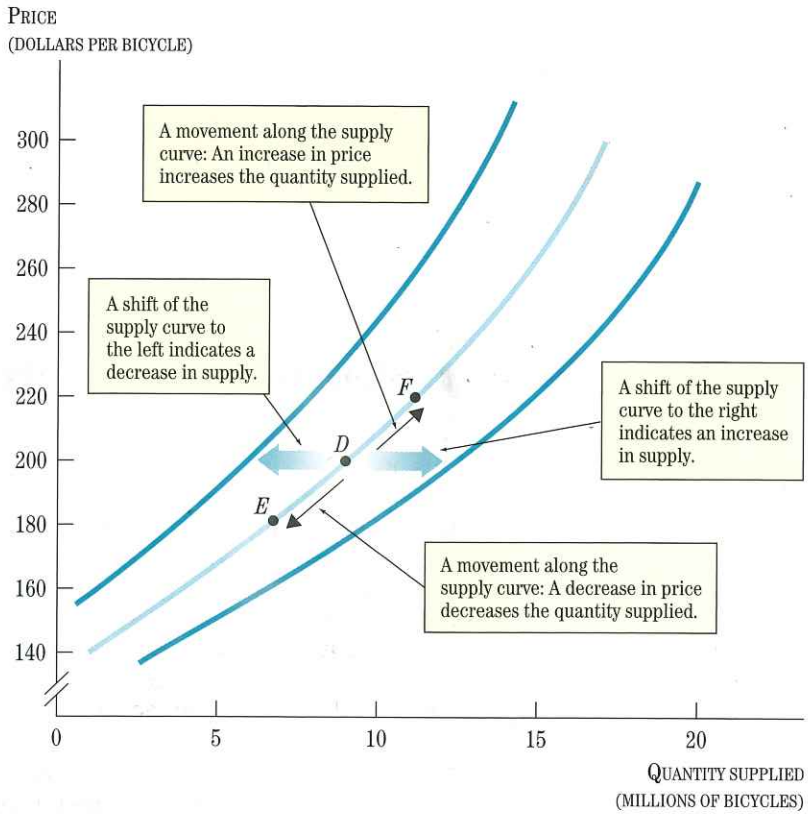


Figure 3.6
Shifts of versus Movements Along the Supply Curve
A *shift* of the supply curve occurs when there is a change in something (other than the price) that affects the amount of a good that firms are willing to supply. An increase in supply is a shift to the right of the supply curve. A decrease in supply is a shift to the left of the supply curve. A movement along the supply curve occurs when the price of the good changes, causing the quantity supplied by firms to change—for example, from point *D* to point *E* or *F*.

The term *supply* refers to the entire supply curve. The term *quantity supplied* refers to a point on the supply curve. As we will soon see, it is important to be able to tell whether a change in something causes (1) a change in supply or (2) a change in the quantity supplied; or, equivalently, (1) a shift in the supply curve or (2) a movement along the supply curve.

Here's an example to test your ability to distinguish between movement along a supply curve and a shift in the supply curve. You will recall that U.S. agricultural policy guarantees farmers a specific price on certain crops. The U.S. government pays these farmers if the price their crops sell for is too low. An economist suggested that the government should instead pay farmers to not plant some of their fields. Which policy is describing a *change in supply* and which is describing a *change in the quantity supplied* in the market for corn?

A policy that pay farmers to leave cornfields unplanted describes a decrease in supply. The amount of corn supplied will be lower at any price. When the U.S. government guarantees the price of corn, this describes an increase in the quantity supplied—more corn will be grown in anticipation of the higher price. The increase in price leading to an increase in quantity supplied corresponds to movement along the supply curve. You can imagine that these policies would lead to differences in the amount of corn grown in the United States and in the price of corn.

REVIEW

- Supply is a positive relationship between the price of a good and the quantity supplied of the good by firms. The supply curve slopes upward because higher prices give firms more incentive to produce and sell more.
- When the quantity supplied changes because of a change in price, we have a movement along the supply curve. Other factors—such as technology, the number of firms, and expectations—affect supply. When these determinants change, the supply curve shifts.

Market Equilibrium: Combining Supply and Demand

Thus far, as summarized in Figure 3.7, we have examined consumers' demand for goods in a market and firms' supply of goods in a market. Now we put supply and demand together to complete the supply and demand model. When consumers buy goods and firms sell goods, they interact in a market, and a price is determined. Recall that a market does not need to be located at one place; the U.S. bicycle market consists of all the bicycle firms that sell bicycles and all the consumers who buy bicycles.

Although it may sound amazing, no single person or firm determines the price in the market. Instead, the market determines the price. As buyers and sellers interact, prices may go up for a while and then go down. Alfred Marshall, the economist who did the most to develop the supply and demand model in the late nineteenth century, called this process the "higgling and bargaining" of the market. The assumption underlying the supply and demand model is that, in the give and take of the marketplace, prices adjust until they settle down at a level where the quantity supplied by firms equals the quantity demanded by consumers. Let's see how.

Table 3.3
Finding the Market Equilibrium

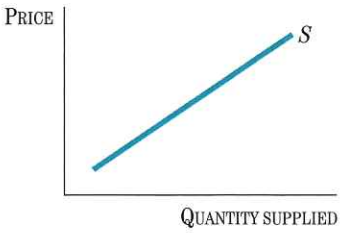
Price	Quantity Demanded	Quantity Supplied	Shortage, Surplus, or Equilibrium	Price Rises or Falls
$140	18	1	Shortage = 17	Price rises
$160	14	4	Shortage = 10	Price rises
$180	11	7	Shortage = 4	Price rises
$200	9	9	Equilibrium	No change
$220	7	11	Surplus = 4	Price falls
$240	5	13	Surplus = 8	Price falls
$260	3	15	Surplus = 12	Price falls
$280	2	16	Surplus = 14	Price falls
$300	1	17	Surplus = 16	Price falls

Quantity supplied equals quantity demanded.

SUPPLY

Supply describes firms.

The supply curve looks like this:



Law of Supply

Price and quantity supplied are positively related.

Movements along supply curve occur

when price rises and quantity supplied rises
or
when price falls and quantity supplied falls.

Shifts in supply are due to:

Technology (new inventions)

Number of firms in market

Price of goods used in production (inputs such as fertilizer, labor)

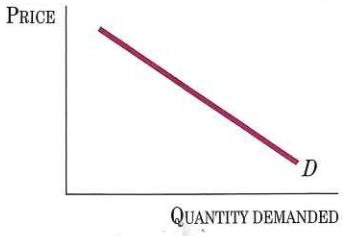
Expectations of future prices (firms will sell less now if prices are expected to rise; for example, farmers may store goods to sell next year)

Government taxes, subsidies, regulations (commodity taxes, agricultural subsidies, safety regulations)

DEMAND

Demand describes consumers.

The demand curve looks like this:



Law of Demand

Price and quantity demanded are negatively related.

Movements along demand curve occur

when price rises and quantity demanded falls
or
when price falls and quantity demanded rises.

Shifts in demand are due to:

Preferences (nice weather or fitness craze changes tastes)

Number of consumers in market

Consumers' information (about smoking, or faulty products, for example)

Consumers' income (normal goods versus inferior goods)

Expectations of future prices (consumers will buy more now if prices are expected to rise in the future)

Price of related goods (both substitutes, like butter and margarine, and complements, like gasoline and SUVs)

Figure 3.7
Overview of Supply and Demand

Determination of the Market Price

To determine the market price, we combine the demand relationship between the price and the quantity demanded with the supply relationship between the price and the quantity supplied. We can do this using either a table or a diagram. First consider Table 3.3, which combines the demand schedule from Table 3.1 with the supply schedule from Table 3.2. The price is in the first column, the quantity demanded by consumers is in the second column, and the quantity supplied by firms is in the third column. Observe that the quantity consumers are willing to buy is shown to decline

with the price, while the quantity firms are willing to sell is shown to increase with the price. In order to determine the price in the market, consider each of the prices in Table 3.3.

■ **Finding the Market Price.** Pick a price in Table 3.3, any price. Suppose the price you choose is \$160. Then the quantity demanded by consumers (14 million bicycles) is greater than the quantity supplied by firms (4 million bicycles). In other words, there is a shortage of $14 - 4 = 10$ million bicycles. A **shortage,** or **excess demand,** is a situation in which the quantity demanded is greater than the quantity supplied. With a shortage of bicycles, the price will quickly rise above \$160; firms will charge higher prices, and consumers who are willing to pay more than \$160 for a bicycle will pay higher prices to firms. Thus, \$160 cannot last as the market price. Observe that as the price rises above \$160, the quantity demanded falls and the quantity supplied rises. Thus, as the price rises, the shortage begins to decrease. If you choose any price below \$200, the same thing will happen: There will be a shortage, and the price will rise. The shortage disappears only when the price rises to \$200, as shown in Table 3.3.

shortage (excess demand): the situation in which quantity demanded is greater than quantity supplied.

Now pick a price above \$200. Suppose you pick \$260. Then the quantity demanded by consumers (3 million bicycles) is less than the quantity supplied by firms (15 million bicycles). In other words, there is a surplus of 12 million bicycles. A **surplus,** or **excess supply,** is a situation in which the quantity supplied is greater than the quantity demanded. With a surplus of bicycles, the price will fall: Firms that are willing to sell bicycles for less than \$260 will offer to sell to consumers at lower prices. Thus, \$260 cannot be the market price either. Observe that as the price falls below \$260, the quantity demanded rises and the quantity supplied falls. Thus, the surplus decreases. If you choose any price above \$200, the same thing will happen: There will be a surplus, and the price will fall. The surplus disappears only when the price falls to \$200.

surplus (excess supply): the situation in which quantity supplied is greater than quantity demanded.

Thus, we have shown that for any price below \$200, there is a shortage, and the price rises; while for any price above \$200, there is a surplus, and the price falls. What if the market price is \$200? Then the quantity supplied equals the quantity demanded; there is neither a shortage nor a surplus, and there is no reason for the price to rise or fall. This price of \$200 is therefore the most likely market price. It is called the **equilibrium price** because at this price the quantity supplied equals the quantity demanded, and there is no tendency for the price to change. There is no other price for which quantity supplied equals quantity demanded. If you look at all the other prices, you will see that there is either a shortage or a surplus, and thus there is a tendency for the price to either rise or fall.

equilibrium price: the price at which quantity supplied equals quantity demanded.

The quantity bought and sold at the equilibrium price is 9 million bicycles. This is the **equilibrium quantity.** When the price equals the equilibrium price and the quantity bought and sold equals the equilibrium quantity, we say that there is a **market equilibrium.**

equilibrium quantity: the quantity traded at the equilibrium price.

market equilibrium: the situation in which the price is equal to the equilibrium price and the quantity traded equals the equilibrium quantity.

Our discussion of the determination of the equilibrium price shows how the market price coordinates the buying and selling decisions of many firms and consumers. We see that the price serves a *rationing function.* That is, the price alleviates shortages: A higher price reduces the quantity demanded or increases the quantity supplied when necessary to eliminate a shortage. Similarly, a lower price increases the quantity demanded or decreases the quantity supplied when there is a surplus. Thus, both shortages and surpluses are eliminated by the forces of supply and demand.

■ **Two Predictions.** By combining supply and demand, we have completed the supply and demand model. The model can be applied to many markets, not just the example of the bicycle market. One prediction of the supply and demand model is that *the price in the market will be the price for which the quantity supplied equals the quantity demanded.* Thus, the model provides an answer to the question of what

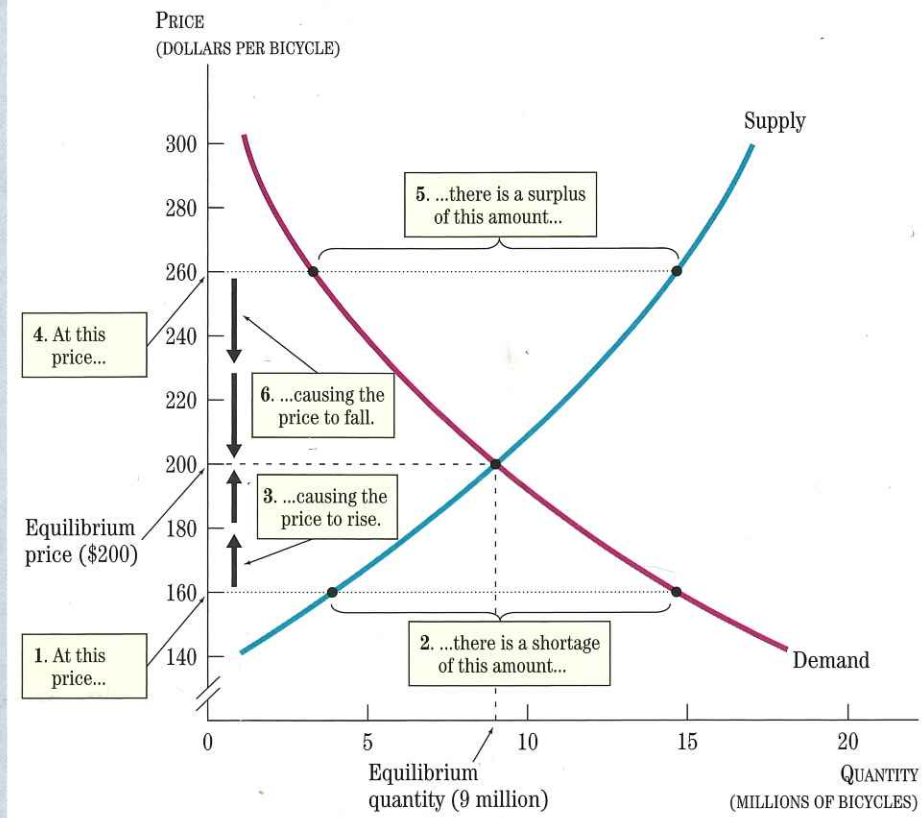


Figure 3.8
Equilibrium Price and Equilibrium Quantity
When buyers and sellers interact in the market, the equilibrium price is at the point of intersection of the supply curve and the demand curve. At this point, the quantity supplied equals the quantity demanded. The equilibrium quantity is also determined at that point. At a higher price, the quantity demanded will be less than the quantity supplied; there will be a surplus. At a lower price, the quantity demanded will be greater than the quantity supplied; there will be a shortage.

determines the price in the market. Another prediction of the model is that *the quantity bought and sold in the market is the quantity for which the quantity supplied equals the quantity demanded.*

Finding the Equilibrium with a Supply and Demand Diagram

The equilibrium price and quantity in a market can also be found with the help of a graph. Figure 3.8 combines the demand curve from Figure 3.1 and the supply curve from Figure 3.4 in the same diagram. Observe that the downward-sloping demand curve intersects the upward-sloping supply curve at a single point. At that point of intersection, the quantity supplied equals the quantity demanded. Hence, the *equilibrium price occurs at the intersection of the supply curve and the demand curve.* The equilibrium price of $200 is shown in Figure 3.8. At that price, the quantity demanded is 9 million bicycles, and the quantity supplied is 9 million bicycles. This is the equilibrium quantity.

If the price were lower than this equilibrium price, say, $160, then the quantity demanded would be greater than the quantity supplied. There would be a shortage, and the price would begin to rise, as shown in the graph. The increase in gasoline prices in 2004 and 2005 led to an increase in demand for hybrid automobiles. With a

shortage of hybrid vehicles and long waiting lists, some automobile sellers increased the price of the hybrids. When there is a shortage, the quantity demanded is greater than the quantity supplied, and there is pressure on the price to increase.

On the other hand, if the price were above the equilibrium price, say, $260, then there would be a surplus, as shown in the graph, and the price would begin to fall. After September 11, 2001, a large number of vacationers cancelled vacation plans that involved air travel. Caribbean hotels facing this decrease in demand began to offer big discounts. When there is a surplus, the quantity supplied is greater than the quantity demanded, and there is pressure on the price to fall.

Thus, the market price will tend to move toward the equilibrium price at the intersection of the supply curve and the demand curve. We can calculate exactly what the equilibrium price is on Figure 3.8 by drawing a line over to the vertical axis. And we can calculate the equilibrium quantity by drawing a line down to the horizontal axis.

A Change in the Market

In order to use the supply and demand model to explain or predict changes in prices, we need to consider what happens to the equilibrium price when there is a change in supply or demand. We first consider a change in demand and then a change in supply.

■ **Effects of a Change in Demand.** Figure 3.9 shows the effects of a shift in the demand curve for bicycles. Suppose that a shift occurs because of a fitness craze that increases the demand for bicycles. The demand curve shifts to the right, as shown in graph (a) in Figure 3.9. The demand curve before the shift and the demand curve after the shift are labeled the "old demand curve" and the "new demand curve," respectively.

If you look at the graph, you can see that something must happen to the equilibrium price when the demand curve shifts. The equilibrium price is determined at the intersection of the supply curve and the demand curve. With the new demand curve, there is a new intersection and, therefore, a new equilibrium price. The equilibrium price is no longer $200 in Figure 3.9(a); it is up to $220 per bicycle. Thus, the supply and demand model predicts that the price in the market will rise if there is an increase in demand. Note also that there is a change in the equilibrium quantity of bicycles. The quantity of bicycles sold and bought has increased from 9 million to 11 million. Thus, the equilibrium quantity has increased along with the equilibrium price. The supply and demand model predicts that an increase in demand will raise both the price and the quantity sold in the market.

We can use the same method to find out what happens if demand decreases, as shown in graph (b) in Figure 3.9. In this case, the demand curve shifts to the left. At the new intersection of the supply and demand curves, the equilibrium price is lower, and the quantity sold is also lower. Thus, the supply and demand model predicts that a decrease in demand will both lower the price and lower the quantity sold in the market.

Note in these examples that when the demand curve shifts, it leads to a movement along the supply curve. First, the demand curve shifts to the right or to the left. Then there is movement along the supply curve because the change in the price affects the quantity of bicycles firms will sell.

■ **Effects of a Change in Supply.** Figure 3.10 shows what happens when there is a change in the market that shifts the supply curve. In graph (a) of Figure 3.10 we

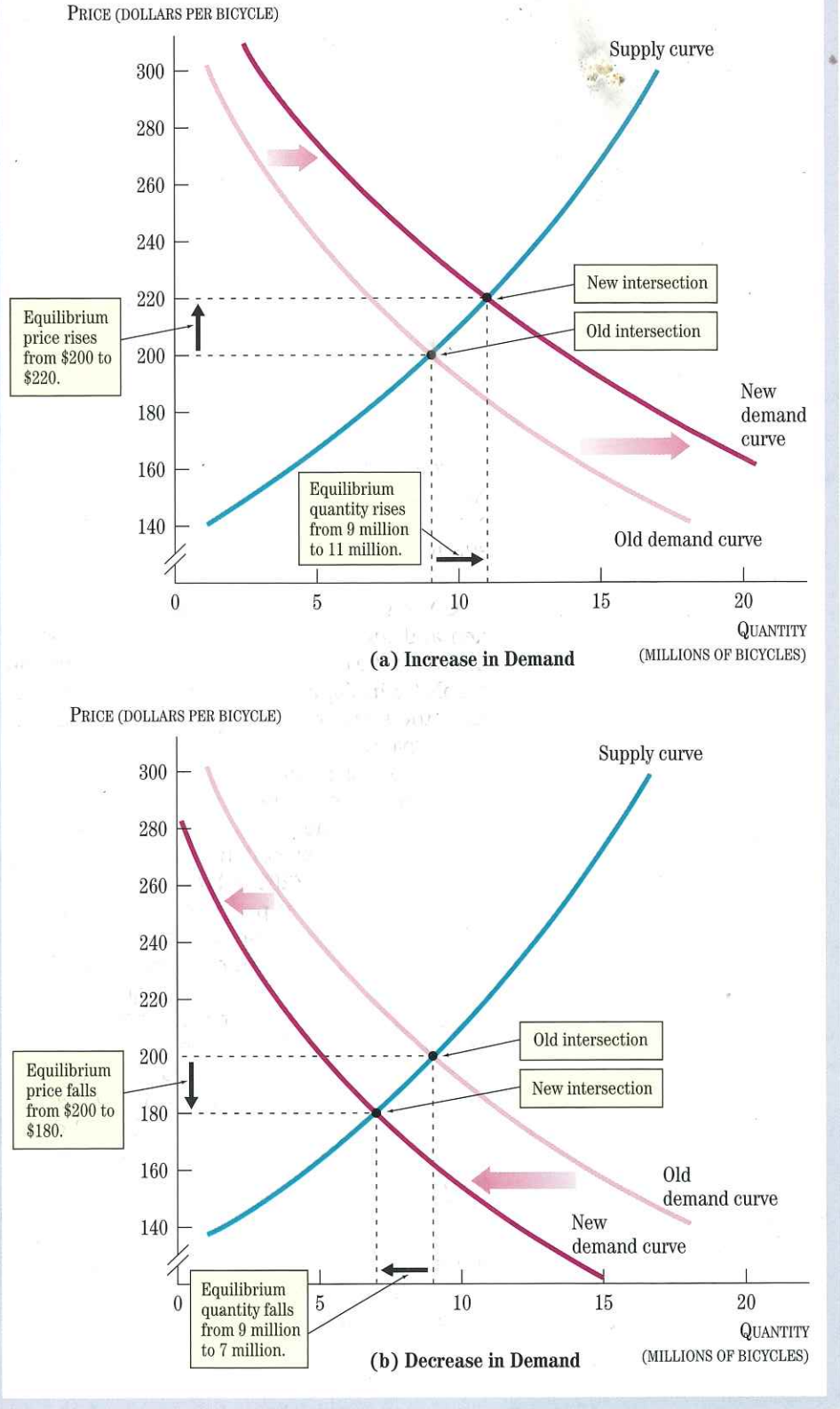
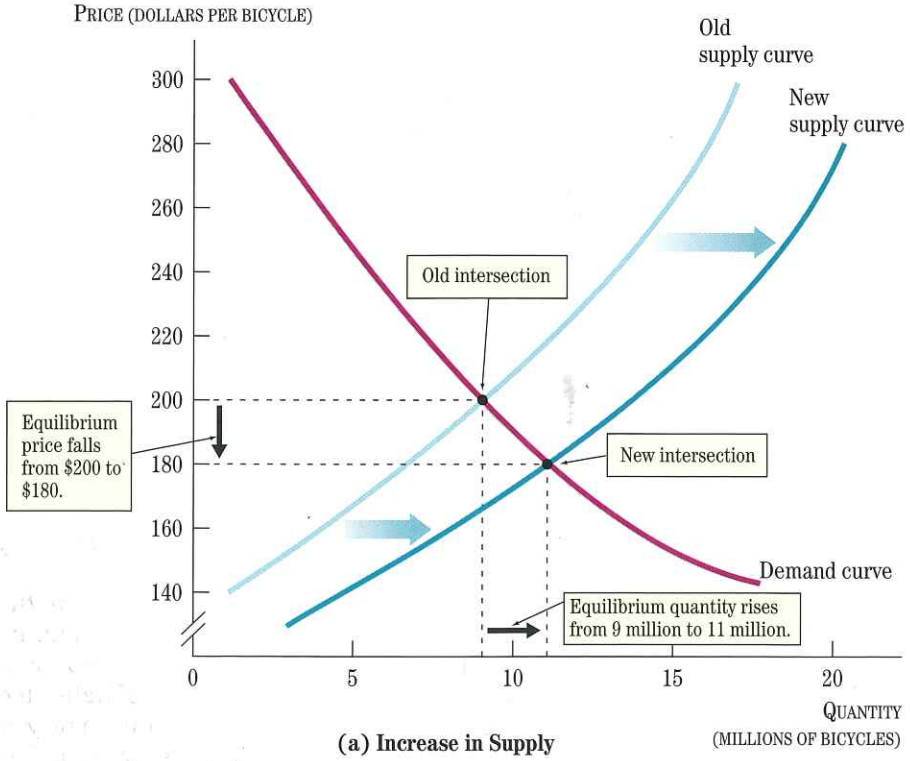
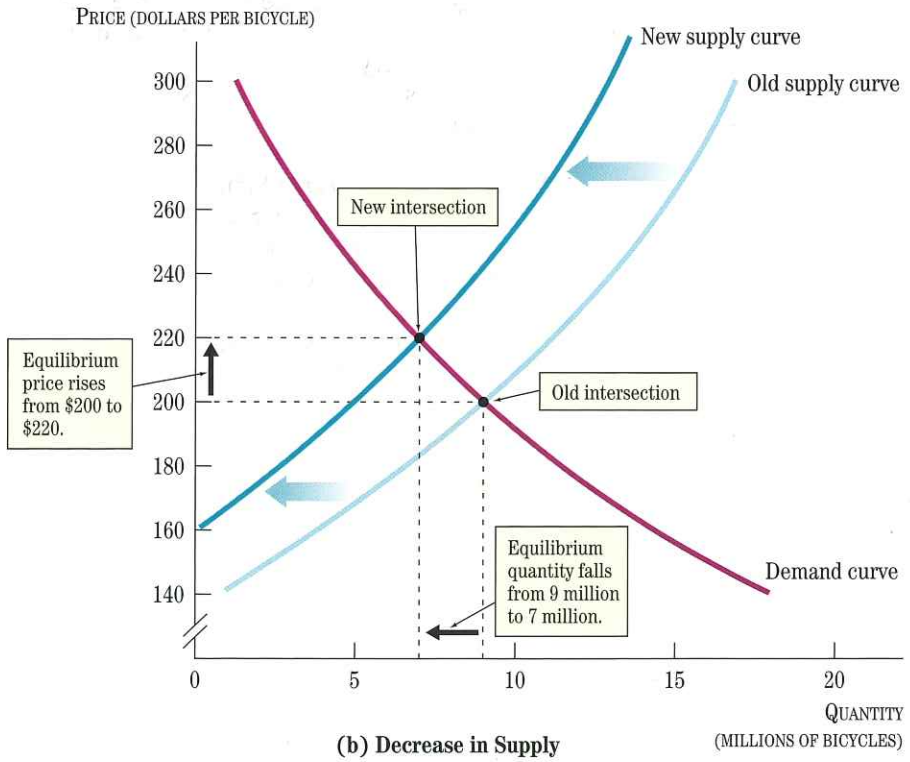


Figure 3.9
Effects of a Shift in Demand
When demand increases, as in graph (a), the demand curve shifts to the right. The equilibrium price rises, and the equilibrium quantity also rises. When demand decreases, as in graph (b), the demand curve shifts to the left. The equilibrium price falls, and the equilibrium quantity also falls.



(a) Increase in Supply



(b) Decrease in Supply

Figure 3.10
Effects of a Shift in Supply
When supply increases, as in graph (a), the supply curve shifts to the right; the equilibrium price falls, and the equilibrium quantity rises. When supply decreases, as in graph (b), the supply curve shifts to the left; the equilibrium price rises, and the equilibrium quantity falls.

Table 3.4
Effects of Shifts in Demand and Supply Curves

Shift	Effect on Equilibrium Price	Effect on Equilibrium Quantity
Increase in demand	Up	Up
Decrease in demand	Down	Down
Increase in supply	Down	Up
Decrease in supply	Up	Down

show the effect of an increase in supply, and in graph (b) we show the effect of a decrease in supply.

When the supply curve of bicycles shifts to the right, there is a new equilibrium price, which is lower than the old equilibrium price. In addition, the equilibrium quantity rises. Thus, the supply and demand model predicts that an increase in the supply of bicycles—perhaps because of better technology in bicycle production—will lower the price and raise the quantity of bicycles sold.

When the supply curve of bicycles shifts to the left, the equilibrium price rises, as shown in graph (b) of Figure 3.10, and the equilibrium quantity falls. Thus, the model predicts that anything that reduces supply will raise the price of bicycles and lower the quantity of bicycles produced.

Table 3.4 summarizes the results of this analysis of shifts in the supply and demand curves.

■ **When Both Curves Shift.** The supply and demand model is easiest to use when something shifts either demand or supply but not both. However, in reality, it is possible for something or several different things to simultaneously shift both supply and demand. To predict whether the price or the quantity rises or falls in such cases, we need to know whether demand or supply shifts by a larger amount. Dealing with the possibility of simultaneous shifts in demand and supply curves is important in practice, as we show in the next section.

REVIEW

- When firms and consumers interact in a market, a price is determined by the market.
- The supply and demand model predicts that the price is found at the intersection of the supply and demand curves. This price is called the equilibrium price.
- At this price, the quantity supplied equals the quantity demanded, and there is no tendency for the price to change.
- A shift in the demand curve or the supply curve will change the equilibrium price and the equilibrium quantity. By considering changes in supply or demand, the model can be used to explain or predict price changes.

CASE STUDY

Using the Supply and Demand Model to Analyze Real-World Issues

In April 2005, President Bush held a press conference to discuss issues that he felt were of critical importance to the American people. One of these issues was the high price of gasoline. How does the high price of gasoline affect you? Perhaps when gasoline gets more expensive, you cut back on driving, or maybe you ask your parents for more money to buy gasoline, or perhaps you just have less money to spend on other things because you are spending more on gasoline. President Bush recognized this impact by asserting, "Millions of American families and small businesses are hurting because of higher gasoline prices."

Why does the price of gasoline go up? How does the high price of gasoline affect the American people and American businesses? What can policymakers do to lower the price of gasoline? The model of supply and demand gives us a tool to model the market for gasoline, to examine the causes of the high price, to comment on the impact on the American people and American businesses, and to focus on what policymakers can do to lower the price of gasoline.

In his April 2005 press conference, President Bush explained the reason for the high price of gasoline. "Over the past decade, America's energy consumption has been growing about 40 times faster than our energy production." A closer analysis of

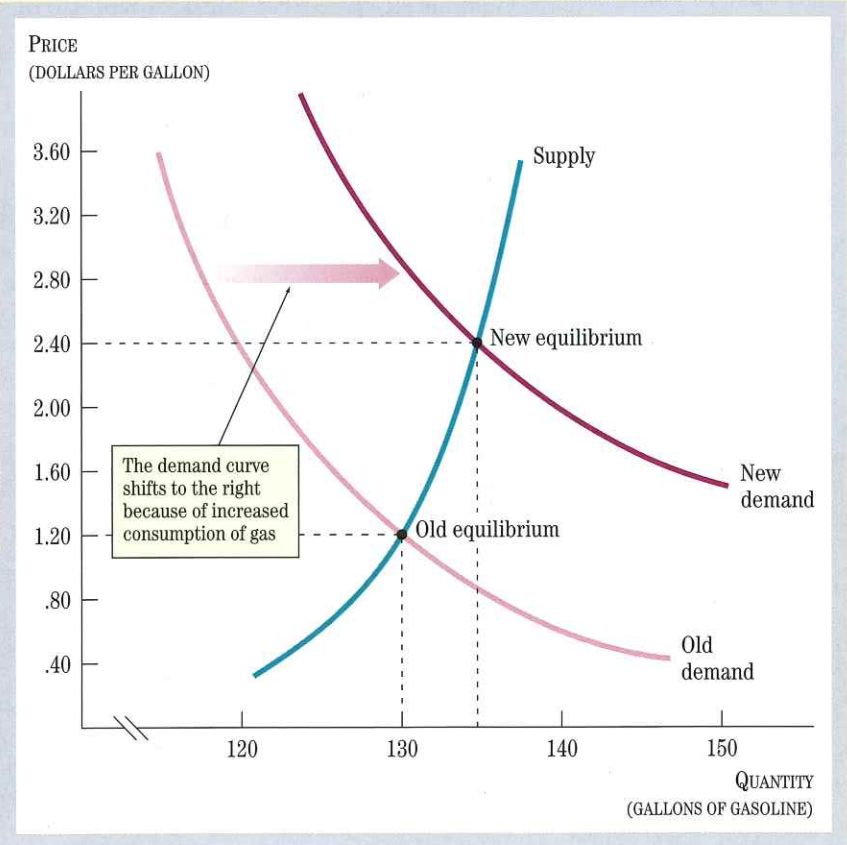


Figure 3.11
Effect of an Increase in Demand on the Supply of Gasoline
You can see here that when demand for gasoline increases, the demand curve shifts to the right. The equilibrium price rises, as does the equilibrium quantity.

the market for gasoline using the model of supply and demand helps us understand the causes of this imbalance in demand and supply. Demand for gasoline has been increasing as more Americans drive gas-guzzling SUVs and Americans drive more miles. Supply of gasoline has been decreasing as a result of a reduction in U.S. refining capacity. Stricter environmental regulations for refining gasoline and the increasing price of oil both lead to an increase in production costs for gasoline. The supply of gasoline decreases as a result of an increase in production costs—that is, a higher price is needed to sell any quantity.

We use a supply and demand model to analyze these changes. As shown in Figure 3.11, increased consumption causes a shift of the demand curve to the right. The quantity demanded is greater than the quantity supplied at the current price, causing the price to rise. When there is an increase in demand and at the same time a decrease in supply, as shown in Figure 3.12, the equilibrium price rises again.

How does the high price of gasoline hurt American businesses and the American people? Here are a few examples. When Wal-Mart's earnings were lower than expected in May 2005, Wal-Mart blamed the low earnings partially on high gasoline prices. Wal-Mart speculated that its customers had less money to spend because of the high price of gasoline. With SUV sales falling in 2005, American auto producers lowered SUV prices to increase their sales. The decrease in sales of SUVs was attributed partially to the high price of gasoline. Think about how your behavior changes when the price of gasoline increases.

What policymakers do to lower the price of gasoline? President Bush stressed in his press conference that Congress needs to pass an energy bill to address the high price of energy. President Bush stated, "You can't wave a magic wand. I wish I could."

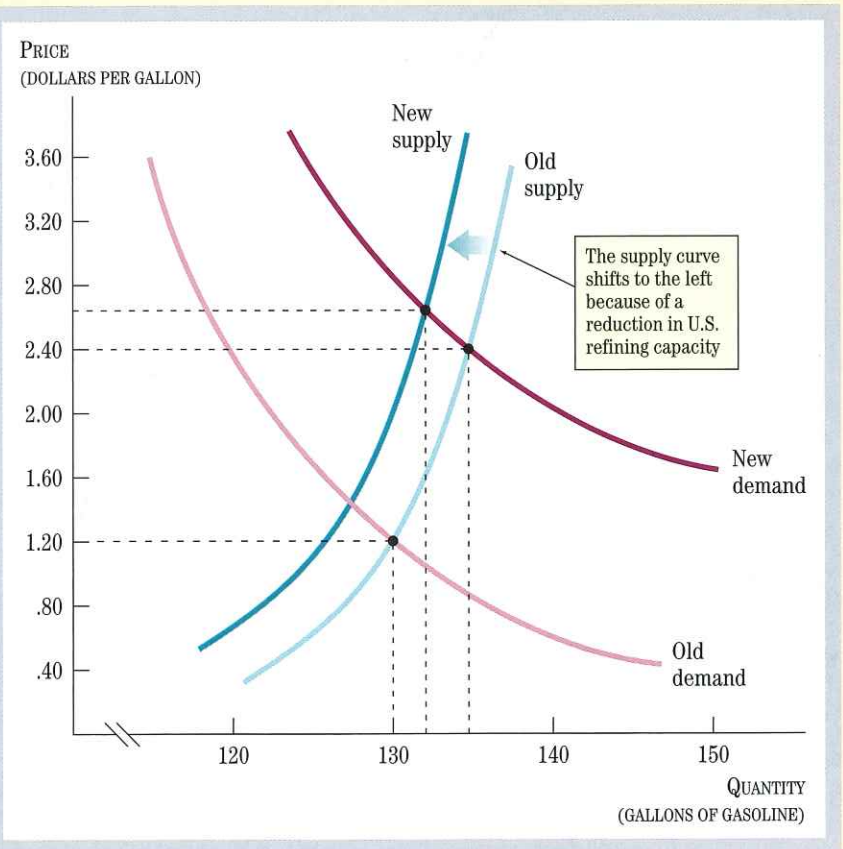


Figure 3.12
Combined Effect of a Simultaneous Increase in Demand and Decrease in Supply of Gasoline
When demand for gasoline increases and, at the same time, the supply of gasoline *decreases* because of decreased refining capacity, the supply curve will shift to the left; again, the equilibrium price increases, but this time the equilibrium quantity decreases, although not back to the original quantity. In this situation, the increase in demand is larger than the decrease in supply.

A magic wand won't work, but the model of supply and demand can predict what will. We'll discuss some of the policies proposed for the energy bill and see that these policies can be modeled as either a decrease in the demand for gasoline, an increase in the supply of gasoline, a decrease in the demand for oil, or an increase in the supply of oil.

President Bush described the essential ingredients of a successful energy policy—it must use new technologies for energy conservation; it must promote the development of new sources of energy; and it must help China and India, both countries with rapidly increasing demand for energy, conserve energy. New technologies for conservation of energy and developing new sources of energy would reduce the demand for oil and for gasoline. This decrease in demand for oil would lead to a decrease in the equilibrium price of oil. Since oil is a major input in the production of gasoline, a decrease in the price of oil corresponds to an increase in the supply of gasoline. Both the decrease in the demand for gasoline and the increase in the supply of gasoline lead to a reduction in the equilibrium price of gasoline. Indeed, this energy policy would help relieve the distress to the American people and businesses caused by the high price of gasoline.

An energy policy that focuses on new technologies for conservation and new sources of energy cannot lead to a quick change in the price of gasoline. President Bush pointed this out, noting that the best way to get the price of gasoline to fall soon would be to encourage oil-producing nations to increase their supply of oil. An increase in the supply of oil would decrease the price of oil, leading to an increase in the supply of gasoline and a reduction in the price of gasoline.

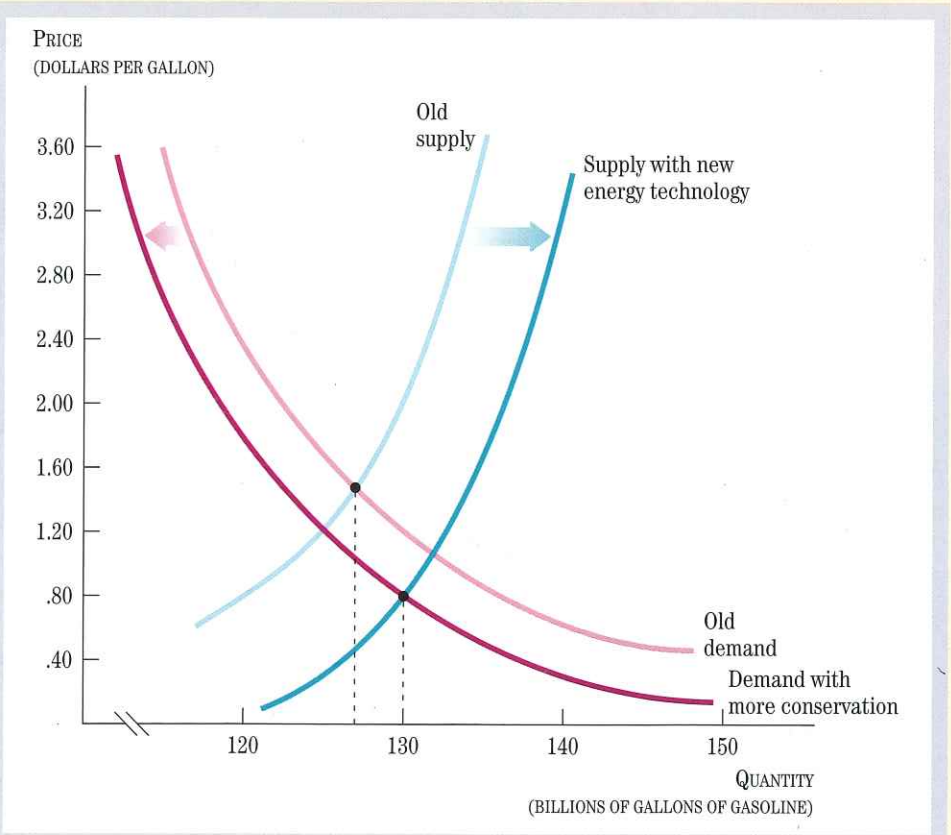


Figure 3.13
Predicted Effects of Energy Policy
The supply and demand model can also be used to predict what would happen with a successful energy policy that promoted the development of new sources of energy and energy conservation. Here, demand decreases slightly due to the effects of energy conservation, and supply increases due to the development of new technology for energy development. When demand decreases and supply increases, the equilibrium price goes down and the equilibrium quantity increases.

Figure 3.13 illustrates the gasoline market with a simultaneous decrease in demand and increase in supply. Both the decrease in the demand for gasoline and the increase in the supply of gasoline would lead to a decrease in the equilibrium price of gasoline. This is a prediction that policymakers could easily make. What if they wanted to also predict the change in the consumption of gasoline resulting from this energy bill? A decrease in demand would decrease equilibrium consumption, while an increase in supply would increase equilibrium consumption. Policymakers therefore could not predict whether gasoline consumption would rise or fall without knowing whether demand or supply would shift by a larger amount. Figure 3.13 shows a resulting increase in the consumption of gasoline because the supply increase is greater in magnitude than the demand decrease. Draw a graph yourself, but make the demand decrease larger than the supply increase. You will see a resulting decrease in the consumption of gasoline. If policymakers want an energy bill that reduces the price of gasoline and reduces the quantity of gasoline consumed, they need to be sure that conservation efforts are the primary focus of the plan.

REVIEW

- The supply and demand model can be used in practical applications to explain price changes in many markets. It can also be used to predict what will happen to prices when certain actions—such as introducing a new energy policy—are taken.
- In applying the model, economists consider shifts of the supply curve or the demand curve. In the case study of gasoline prices, both the supply curve and the demand curve shifted.

Interference with Market Prices

Thus far, we have used the supply and demand model in situations in which the price is freely determined without government control. But many times throughout history, and around the world today, governments have attempted to control market prices. The usual reason is that government leaders have not been happy with the outcome of the market, or they were pressured by groups who would benefit from price controls.

price control: a government law or regulation that sets or limits the price to be charged for a particular good.

Price controls were used widely by the U.S. government during World War II and again in the early 1970s. Price controls now exist in certain housing markets, agriculture markets, and labor markets in the United States. What are the effects of this government interference in the market? The supply and demand model can help answer this question.

Price Ceilings and Price Floors

price ceiling: a government price control that sets the maximum allowable price for a good.

In general there are two broad types of government price controls. Controls can stipulate a **price ceiling,** or a maximum price at which a good can be bought and sold. For example, the United States government controlled oil prices in the early 1970s, stipulating that firms could not charge more than a stated maximum price of $5.25 per barrel of crude oil; the equilibrium price was well over $10 per barrel at this time. Some cities in the United States have price controls on rental apartments. Landlords

rent control: a government price control that sets the maximum allowable rent on a house or apartment.

price floor: a government price control that sets the minimum allowable price for a good.

minimum wage: a wage per hour below which it is illegal to pay workers.

are not permitted to charge a rent higher than the maximum stipulated by the **rent control** law in these cities. Price ceilings are imposed by governments because of complaints that the market price is too high. The purpose is to help the consumers who must pay the prices. For example, rent controls exist in order to help people who must pay rent. However, as we will see, price controls have harmful side effects that can end up hurting those consumers the law is apparently trying to help.

Government price controls can also stipulate a **price floor,** or a minimum price. Price floors are imposed by governments in order to help the suppliers of goods and services. For example, the U.S. government requires that the price of sugar not fall below a certain amount in the United States. In the labor market, the U.S. government requires that firms pay workers a wage of at least a given level, called the **minimum wage.**

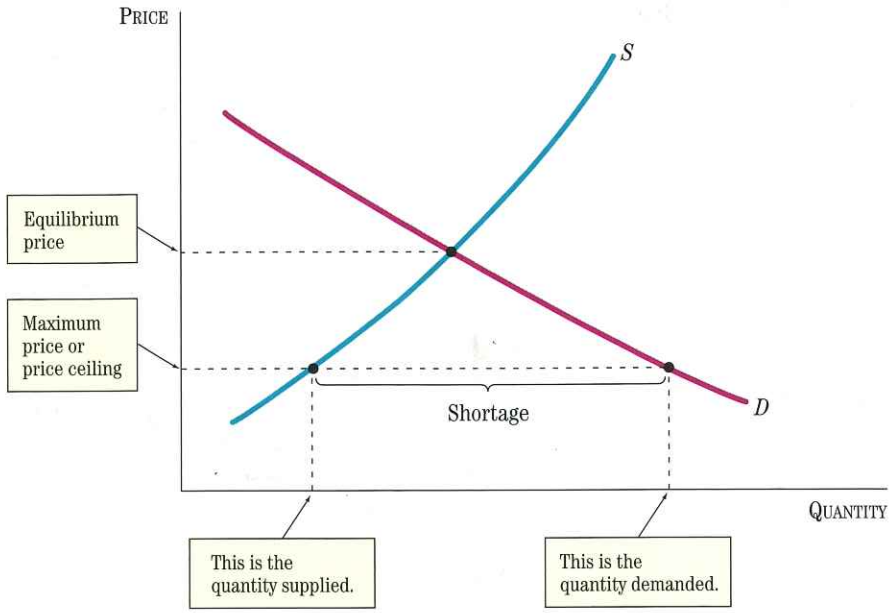
Shortages and Related Problems Resulting from Price Ceilings

If the government prevents firms from charging more than a certain amount for their products, then a shortage is likely to result, as illustrated in Figure 3.14. When the maximum price remains below the equilibrium price for the market, there is a persistent shortage; sellers are unwilling to supply as much as buyers want to buy. This is illustrated for the general case of any good in the top graph in Figure 3.14 and for the specific case of rent control in the bottom graph.

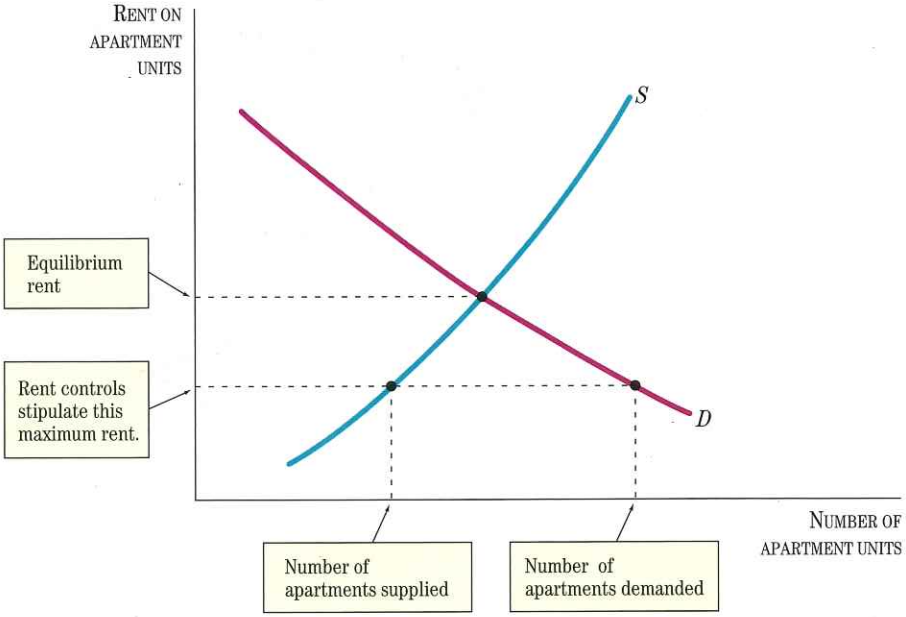
■ **Dealing with Persistent Shortages.** Because higher prices are not allowed, the shortage must be dealt with in other ways. Sometimes the government issues a limited amount of ration coupons to people to alleviate the shortage; this was done in World War II. The law required that people present these ration coupons at stores in order to buy goods. Thus, the total quantity demanded could not be greater than the amount of ration coupons. Alternatively, if there are no ration coupons, then the shortage might result in long waiting lines. In the past, in centrally planned economies, long lines for bread were frequently observed because of price controls on bread. Sometimes black markets develop, in which people buy and sell goods outside the watch of the government and charge whatever price they want. This typically happens in command economies. Black markets are also common in less-developed countries today when the governments in these countries impose price controls.

Another effect of price ceilings is a reduction in the quality of the good sold. By lowering the quality of the good, the producer can reduce the costs of producing it. Low-quality housing frequently results from rent control. By lowering the quality of the apartments—perhaps being slow to paint the walls or repair the elevator—landlords make the apartments shoddy and unattractive.

■ **Making Things Worse.** Although the stated purpose of price ceilings is to help people who have to pay high prices, the preceding examples indicate how they can make things worse. Issuing ration coupons raises difficult problems about who gets the coupons. In the case of a price ceiling on gasoline, for example, should the government give more coupons to those who commute by car than to those who do not? More generally, who is to decide who deserves the coupons? Rationing by waiting in line is also a poor outcome. People waiting in line could be doing more enjoyable or more useful things. Similarly, black markets, being illegal, encourage people to go



The General Case of a Price Ceiling



The Case of Rent Control

Figure 3.14
Effects of a Maximum Price Law
The top diagram shows the general case when the government prevents the market price from rising above a particular maximum price, or sets a price ceiling below the equilibrium price. The lower diagram shows a particular example of a price ceiling, rent controls on apartment units. The supply and demand model predicts that there will be a shortage. The shortage occurs because the quantity supplied is less than consumers are willing to buy at that price. The shortage leads to rationing, black markets, or lower product quality.

outside the law; people thereby lose their rights to protection in the case of theft or fraud. Lowering the quality of the good is also a bad way to alleviate the problem of a high price. This simply eliminates the higher-quality good from production; consumers and producers lose.

Paradoxically, price ceilings frequently end up hurting those they try to help. Many people who benefit from controls, for example, are not poor at all. If rent controls reduce the supply of apartments, they make less housing available for everyone.

Surpluses and Related Problems Resulting from Price Floors

If the government imposes a price floor, then a surplus will occur, as shown in Figure 3.15. With the price above the equilibrium price, suppliers of goods and services want to sell more than people are willing to buy. Hence, there is a surplus. This is illustrated for the general case of any good in the top graph of Figure 3.15 and for the specific case of the minimum wage in the bottom graph.

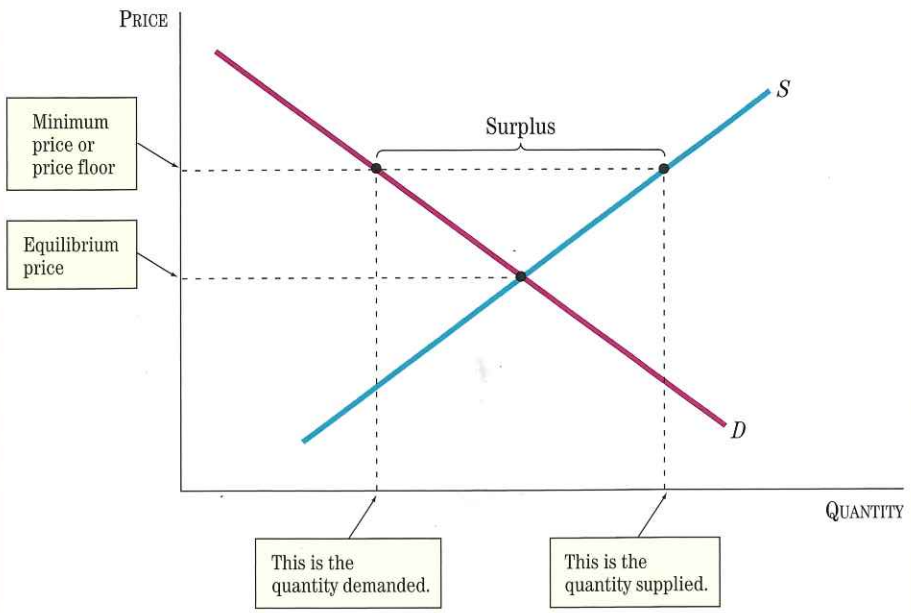
How is this surplus dealt with in actual markets? In markets for farm products, the government usually has to buy the surplus and, perhaps, put it in storage; but buying farm products above the equilibrium price costs taxpayers money, and the higher price raises costs to consumers. For this reason, economists argue against price floors on agricultural goods. As an alternative, the government sometimes reduces the supply by telling firms to plant fewer acres or to destroy crops, or by restricting the amount that can come from abroad. In the United States, the federal government uses acreage restrictions in the case of wheat and other grains; it also uses import restrictions in the case of sugar. But government requirements that land be kept idle or crops destroyed are particularly repugnant to most people.

As we will see in more detail later in this book, the supply and demand model can also be applied to labor markets. In that case, the price is the price of labor, or the wage. What does the supply and demand model predict about the effects of a minimum wage? In the case of labor markets, a minimum wage can cause unemployment. If the equilibrium wage is below the minimum wage, then some workers would be willing to work for less than the minimum wage. But employers are not permitted to pay them less than the minimum wage. Therefore, there is an oversupply of workers at the minimum wage. The number of workers demanded is less than the number of workers willing to work; thus, the supply and demand model predicts that the minimum wage causes unemployment.

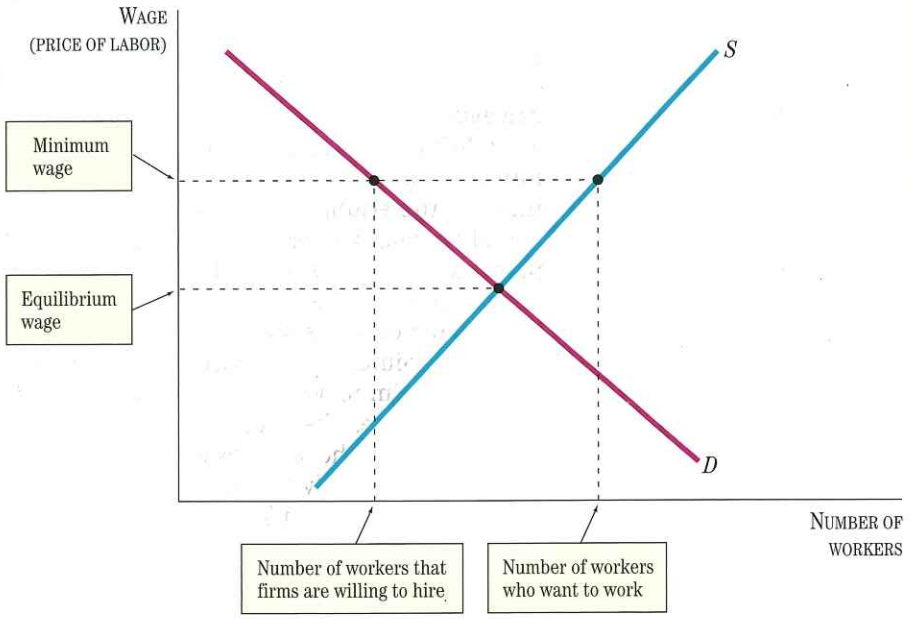
The minimum wage would have no effect if the equilibrium wage were above the minimum wage. The supply and demand model predicts that the minimum wage affects workers whose wages would be below the minimum. Thus, a minimum wage would be most likely to increase unemployment for teenage workers with very few skills if their wages would otherwise be below the minimum.

REVIEW

- Price ceilings cause persistent shortages, which, in turn, cause rationing, black markets, and a reduced quality of goods and services.
- Price floors cause persistent surpluses and unemployment, according to the supply and demand model. In the case of price floors on agricultural products, the surpluses are bought by the government and put in storage. In the case of the minimum wage, the surpluses mean more unemployment for those who can get jobs only at a wage below the minimum wage.



The General Case of a Price Floor



The Case of the Minimum Wage

Figure 3.15
Effects of a Minimum Price Law
The top diagram shows the general case when the government prevents the market price from falling below a particular minimum price, or sets a price floor above the equilibrium price. The lower diagram shows a particular example when the price of labor—the wage—cannot fall below the minimum wage. The supply and demand model predicts that sellers are willing to sell a greater quantity than buyers are willing to buy at that price. Thus, there is a surplus of the good or, in the case of labor, unemployment for some of those who would be hired only at a lower wage.

Conclusion

This chapter has shown how prices are determined in markets where buyers and sellers interact freely. The supply and demand model is used to describe how prices are determined in such markets. It is probably the most frequently used model in economics and has been in existence for over a hundred years in pretty much the same form as economists use it now. You will come to appreciate it more and more as you study economics.

A key feature of the model is that the price is found by the intersection of the supply and demand curves. To apply the model in practice, we need to look for factors that shift either the supply curve or the demand curve. In the most successful applications of the supply and demand model, the factors that affect supply and demand can be separated.

In later chapters we will see more about how the supply and demand model is implied by the central economic idea that people make purposeful choices with scarce resources and interact with other people as they make these choices. By doing so, we will be able to take a closer look at the three basic questions: what, how, and for whom to produce.

KEY POINTS

1. Demand is a negative relationship between the price of a good and the quantity demanded by consumers. It can be shown graphically by a downward-sloping demand curve.
2. A movement along the demand curve occurs when a higher price reduces the quantity demanded or a lower price increases the quantity demanded.
3. A shift of the demand curve occurs when something besides the price causes the quantity people are willing to buy to change.
4. Supply is a positive relationship between the price of a good and the quantity supplied by firms. It can be shown graphically by an upward-sloping supply curve.
5. A movement along the supply curve occurs when a higher price increases the quantity supplied or a lower price decreases the quantity supplied.
6. A shift of the supply curve occurs when something besides the price causes the quantity firms are willing to sell to change.
7. The equilibrium price and equilibrium quantity are determined by the intersection of the supply curve and the demand curve, where the quantity supplied equals the quantity demanded.
8. By shifting either the supply curve or the demand curve, observations on prices can be explained and predictions about prices can be made.
9. Price ceilings cause shortages, with the quantity supplied less than the quantity demanded. Shortages lead to rationing or black markets.
10. Price floors cause surpluses, with the quantity supplied greater than the quantity demanded.

KEY TERMS

demand
price
quantity demanded
demand schedule
law of demand
demand curve
normal good
inferior good
substitute
complement
supply
quantity supplied
supply schedule
law of supply
supply curve
shortage (excess demand)
surplus (excess supply)
equilibrium price
equilibrium quantity
market equilibrium
price control
price ceiling
rent control
price floor
minimum wage

QUESTIONS FOR REVIEW

1. Why does the demand curve slope downward?
2. Why does the supply curve slope upward?
3. What is the difference between a shift in the demand curve and a movement along the demand curve?
4. What are four things that cause a demand curve to shift?
5. What are four things that cause a supply curve to shift?
6. What is the difference between a shift in the supply curve and a movement along the supply curve?
7. What are the equilibrium price and equilibrium quantity?
8. What happens to the equilibrium price if the supply curve shifts to the right and the demand curve does not shift?
9. What happens to the equilibrium price if the demand curve shifts to the right and the supply curve does not shift?
10. Do price ceilings cause shortages or surpluses? What about price floors? Explain.

PROBLEMS

1. Consider the following supply and demand model of the world tea market (in billions of pounds).

Price per Pound	Quantity Supplied	Quantity Demanded
$.38	1,500	525
$.37	1,000	600
$.36	700	700
$.35	600	900
$.34	550	1,200

   a. Is there a shortage or a surplus when the price is $.38? What about $.34?
   b. What are the equilibrium price and the equilibrium quantity?
   c. Graph the supply curve and the demand curve.
   d. Show how the equilibrium price and quantity can be found on the graph.
   e. If there is a shortage or surplus at a price of $.38, calculate its size in billions of pounds and show it on the graph.

2. Consider the supply and demand model in problem 1. Suppose that there is a drought in Sri Lanka that reduces the supply of tea by 400 billion pounds at every price. Suppose demand does not change.
   a. Write down in a table the new supply schedule for tea.
   b. Find the new equilibrium price and the new equilibrium quantity. Explain how the market adjusts to the new equilibrium.
   c. Did the equilibrium quantity change by more or less than the change in supply?
   d. Graph the new supply curve along with the old supply curve and the demand curve.
   e. Show the change in the equilibrium price and the equilibrium quantity on the graph.

3. Use the supply and demand model to explain what happens to the equilibrium price and the equilibrium quantity for frozen yogurt in the following cases:
   a. There is a large expansion in the number of firms producing frozen yogurt.
   b. It is widely publicized in the press that frozen yogurt isn't as healthful as was previously thought.
   c. There is a sudden increase in the price of milk, which is used to produce frozen yogurt.
   d. Frozen yogurt suddenly becomes popular because a movie idol promotes it in television commercials.

4. For each of the following markets, indicate whether the stated change causes a shift in the supply curve, a shift in the demand curve, a movement along the supply curve, and/or a movement along the demand curve.
   a. The housing market: Consumers' incomes fall.
   b. The tea market: The price of sugar goes down.
   c. The coffee market: There is a freeze in Brazil that severely damages the coffee crop.
   d. The fast food market: The number of fast food restaurants in an area decreases.
   e. The peanut market in the U.S. Southeast: A drought lowers supply.

5. Draw a supply and demand diagram to indicate the market for prescription drugs in the United States, with the equilibrium price and quantity labeled. Suppose the government imposes a strict ceiling on the price of prescription drugs sold in the United States. Show what happens in this market if the ceiling is less than the equilibrium price. How would the pharmaceutical firms that develop such drugs in their research laboratories respond?

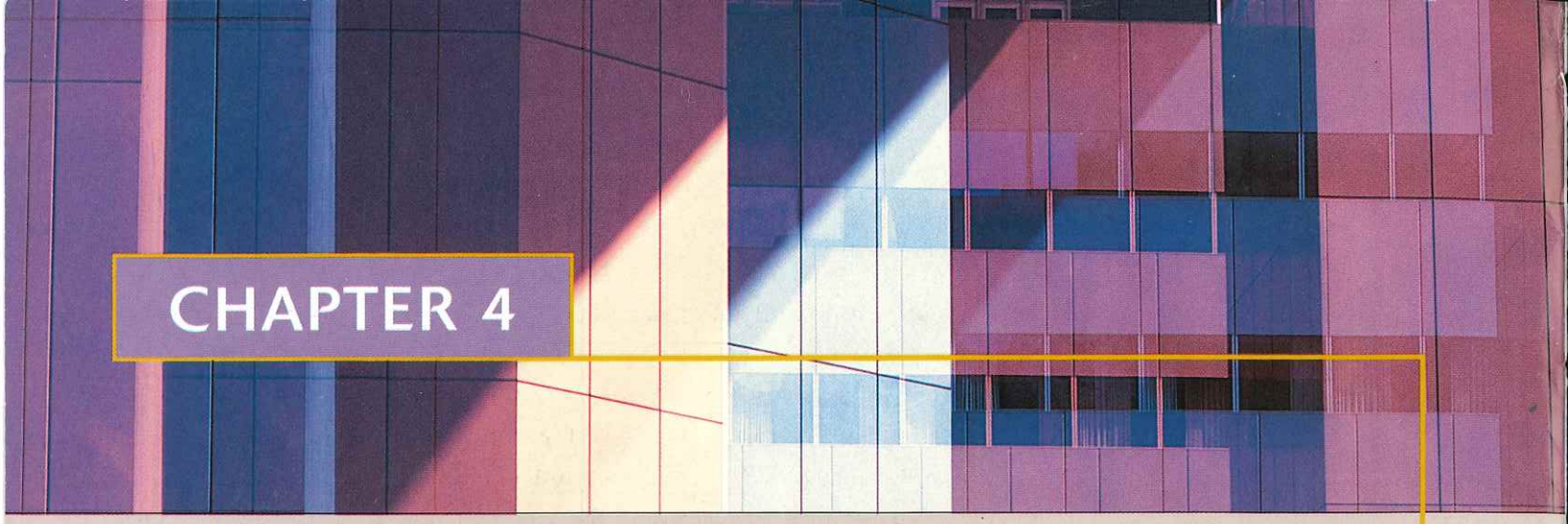
6. Consider the market for automatic teller machine services in a city. The price is the fee for a cash withdrawal.
   a. Sketch the demand curve and the supply curve for ATM transactions.
   b. How is the equilibrium price determined?
   c. If the town council imposes a ban on ATM fees, equivalent to a price ceiling in this market, what happens to quantity supplied and quantity demanded?
   d. Economists frequently argue against price controls because of the incentives they give to suppliers. Explain why this interference in the market may provide bad incentives.

7. In 1991 the price of milk fell 30 percent. Senator Leahy of Vermont, a big milk-producing state, supported a law in the U.S. Congress to put a floor on the price. The floor was $13.09 per hundred pounds of milk. The market price was $11.47.
   a. Draw a supply and demand diagram. Explain the effects of the legislation. Would the legislation cause a surplus or a shortage?
   b. The dairy farmers supported the legislation, and consumer groups opposed it. Why?
8. Why is it necessary for people to stand in line for days before the sale of tickets to concerts by the most famous performers? Is the price mechanism working properly? Why are scalpers present on these occasions?
9. Assuming that either supply or demand, but not both, changes, indicate the direction and change in either supply or demand that must have occurred to produce the following:
   a. A decrease in the price and quantity of apples
   b. A decrease in the price of bananas with an increase in the quantity of bananas
   c. An increase in the price and quantity of cars
   d. An increase in the quantity of computers with a decrease in the price
10. Using the demand and supply diagrams (one for each market), show what short-run changes in price and quantity would be expected in the following markets if worries about air safety cause travelers to shy away from air travel. Each graph should contain the original and new demand and supply curves, and the original and new equilibrium prices and quantities. For each market, write one sentence explaining why each curve shifts or does not shift.
    a. The market for air travel
    b. The market for rail travel
    c. The market for hotel rooms in Hawaii
11. Determine which of the following four sentences use the terminology of the supply and demand model correctly.
    a. "The price of bicycles rose, and therefore the demand for bicycles went down."
    b. "The demand for bicycles increased, and therefore the price went up."
    c. "The price of bicycles fell, decreasing the supply of bicycles."
    d. "The supply of bicycles increased, and therefore the price of bicycles fell."
12. a. Suppose you find out that an increase in the price of first-class postage leads to an increase in the demand for overnight delivery service and a decrease in the demand for envelopes. For which good is postage a complement, and for which is it a substitute?
    b. Suppose someone told you that an increase in the price of gasoline caused a decrease in the demand for public transportation by train. Is this what you would predict? Why or why not?
    c. Suppose an economic forecasting group has determined that an increase in the price of orange juice has no effect on the demand for soft drinks. What can you conclude from this information?
13. Suppose a decrease in consumers' incomes causes a decrease in the demand for chicken and an increase in the demand for potatoes. Which good is inferior and which is normal? How will the equilibrium price and quantity change for each good?
14. a. Straight-line demand and supply curves can be represented by linear algebraic equations. Given the following algebraic expressions for supply and demand, calculate the equilibrium price and quantity by solving the two equations for P and Q.

    Supply: $Q = 5 + 2P$

    Demand: $Q = 9 - 2P$

    b. For the equations defined in part (a), show that when you substitute the equilibrium price into either the supply or the demand equation, you get the same equilibrium quantity.
    c. Suppose that the demand curve shifts as a result of an increase in consumers' incomes. The new demand equation is $Q = 13 - 2P$. Calculate the new equilibrium price and quantity.



CHAPTER 4

Elasticity and Its Uses

When oil prices rose to over $55 a barrel in April 2005—nearly 50 percent higher than the year before—it was "*déja vu* all over again." Oil prices had been rising steadily since 2000 and had also risen sharply several times before—in 1990, in 1980, and in the early 1970s. Each time the story was similar: Oil-producing countries cut back the supply of oil, and the price of oil rose. In 1990 the supply was cut because Iraq invaded Kuwait and destroyed its oil fields. In 2000 the supply was cut when the producing countries got together to raise the price. But this time, the story was somewhat different. There were no critical cuts in supply from any of the major oil-exporting countries. Instead, the world's oil producers were just not able to keep up with the continually increasing demand for oil from around the world, an increase fueled in large part by a significantly increasing demand from China. Because oil is a key input to many products, from the gasoline that people put in their cars and trucks to the asphalt they drive them on, higher oil prices have a serious impact on many people.

The supply and demand model tells us that a reduction in oil supply or an increase in oil demand will increase the price of oil. But by how much? For example, would the price of oil be expected to nearly double when Iraq's military actions removed 7 percent of the world's oil supply? And how much would the price fall if the U.S. government persuaded the oil-producing countries in April 2005 to increase production by 4 percent?

There is an elegant, but remarkably useful, economic concept called *elasticity* that economists use when they work with the supply and demand model. In economics, elasticity is a measure of how sensitive one variable is to another. In the case of the supply and demand model, elasticity measures how sensitive the quantity of a good that people demand, or that firms supply, is to the price of the good. In this chapter we show how the concept of elasticity can be used to answer the questions raised above about oil supply and price and many other questions. We first show why elasticity is important and provide a formula to show how elasticity is calculated. We then show how it is used in many different ways and demonstrate how to work with and talk about elasticity.

Elasticity of Demand

Defining the Price Elasticity of Demand

The price elasticity of demand is a measure of the sensitivity of the *quantity demanded* of a good to the *price* of the good. "Price elasticity of demand" is sometimes shortened to "elasticity of demand," the "demand elasticity," or even simply "elasticity" when the meaning is clear from the context. The price elasticity of demand always refers to a particular demand curve or demand schedule, such as the world demand for oil or the U.S. demand for bicycles. For a particular demand curve, all other things besides the price of the good being equal, these relationships hold: As the price increases, the quantity demanded by consumers declines; as the price decreases, the quantity demanded by consumers increases. The price elasticity of demand is a measure of *how much* the quantity demanded changes when the price changes.

For example, when economists report that the price elasticity of demand for contact lenses is high, they mean that the quantity of contact lenses demanded by people changes by a large amount when the price changes. Or if they report that the price elasticity of demand for bread is low, they mean that the quantity of bread demanded changes by only a small amount when the price of bread changes.

price elasticity of demand: the percentage change in the quantity demanded of a good divided by the percentage change in the price of that good.

We can define the price elasticity of demand clearly with a formula: **Price elasticity of demand** is the percentage change in quantity demanded divided by the percentage change in the price. That is,

$$\text{Price elasticity of demand} = \frac{\text{percentage change in quantity demanded}}{\text{percentage change in the price}}$$

We emphasize that the price elasticity of demand refers to a particular demand curve; thus, the numerator of this formula is the percentage change in quantity demanded when the price changes by the percentage amount shown in the denominator. All the other factors that affect demand are held constant when we compute the price elasticity of demand.

For example, the price elasticity of demand for gasoline is about .2. Thus, if the price of gasoline increases by 10 percent, the quantity of gasoline demanded will fall by 2 percent ($.2 \times 10$). The price elasticity of demand for alcoholic beverages is about 1.5; thus, if the price of alcoholic beverages rises by 10 percent, the quantity

demanded will fall by 15 percent (1.5 × 10). As you can see from these examples, elasticity is a way to determine by how much the *quantity demanded* changes when the price changes.

The Size of the Elasticity: High versus Low

There are two graphs in Figure 4.1, each showing a different possible demand curve for oil in the world. We want to show why it is important to know which of these two demand curves is correct, or at least which one gives a better description of economic behavior in the oil market. Each graph has the price of oil on the vertical axis (in dollars per barrel) and the quantity of oil demanded on the horizontal axis (in millions of barrels of oil a day).

Both of the demand curves pass through the same point *A*, where the price of oil is $20 per barrel and the quantity demanded is 60 million barrels per day. But observe

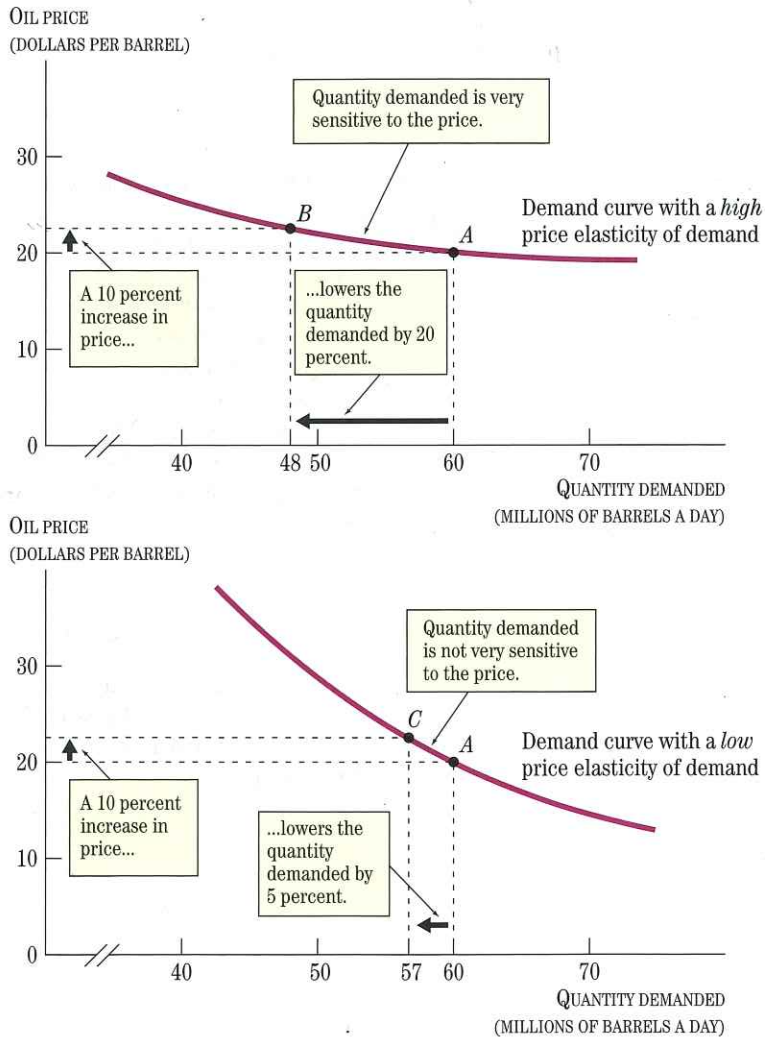


Figure 4.1
Comparing Different Sizes of the Price Elasticity of Demand
Both sets of axes have exactly the same scale. In the top graph, the quantity demanded is very sensitive to the price; the elasticity is high. In the bottom graph, the quantity demanded is not very sensitive to the price; the elasticity is low. Thus, the same increase in price ($2, or 10 percent) reduces the quantity demanded much more when the elasticity is high (top graph) than when it is low (bottom graph).

How Policymakers Use Price Elasticity of Demand to Discourage Underage Drinking

Policymakers use information about the price of elasticity of demand in many ways. Take the government's efforts to reduce underage drinking. In a 2003 study on underage drinking, the National Academy of Sciences recommended that one way to reduce underage drinking would be to increase the tax on alcohol. To implement this policy effectively, it would be important for policymakers to know which demand curve most accurately represents the demand for alcohol by underage drinkers. The amount that a tax would reduce the quantity of alcohol consumed by underage drinkers depends on their price elasticity of demand.

Recall that a new tax is modeled as a decrease in supply. You can see how this works by drawing this supply shift and a demand curve with high price elasticity of demand and then drawing the same supply shift and a demand curve with a low price elasticity of demand, as in Figure 4.2. Alcohol consumption responds more to the tax when the price elasticity of demand is high. If the price elasticity of demand for alcohol by underage drinkers is low (that is, if the quantity of alcohol demanded by underage drinkers changes by only a small amount when the price of alcohol changes), then a new tax on alcohol must be large to accomplish the goal of a reduction in underage drinking. If the price elasticity of demand for alcohol by underage drinkers is high (that is, if the quantity of alcohol demanded by underage drinkers changes by a large amount when the price of alcohol changes), then the tax might not need to be very big to accomplish the policymakers' goal. Which do you think is more likely?

that the two different curves show different degrees of sensitivity of the quantity demanded to the price. In the top graph, where the demand curve is relatively flat, the quantity demanded of oil is very sensitive to the price; in other words, the demand curve has a high elasticity. For example, consider a change from point *A* to point *B:* When the price rises by $2, from $20 to $22, or by 10 percent ($2/$20 = .10, or 10 percent), the quantity demanded falls by 12 million, from 60 million to 48 million barrels a day, or by 20 percent (12/60 = .20 or 20 percent).

On the other hand, in the bottom graph, the quantity demanded is not very sensitive to the price; in other words, the demand curve has a low elasticity. It is relatively steep. When the price rises by $2, or 10 percent, from point *A* to point *C,* the quantity demanded falls by 3 million barrels, or only 5 percent. Thus, the sensitivity of the quantity to the price, or the size of the elasticity, is what distinguishes these two graphs.

The Impact of a Change in Supply on the Price of Oil

Now consider what happens when there is a decline in supply in the world oil market. In Figure 4.2 we combine the supply curve for oil with the two possible demand curves for oil from Figure 4.1. Initially the oil market is in equilibrium in Figure 4.2; in both graphs, the quantity demanded equals the quantity supplied. The equilibrium price is $20 per barrel and the equilibrium quantity is 60 million barrels a day, just like at point *A* in Figure 4.1. A reduction in the supply of oil—perhaps because of the invasion of Kuwait by Iraq in 1990, or perhaps because of the reduction in production by Iraq in 2000—is also shown. The exact same leftward shift in supply is shown in the top and bottom graphs of Figure 4.2.

Now, observe how the equilibrium price changes in the two graphs. Recall that this change is our prediction—using the supply and demand model—of what would happen to the price of oil if the supply declined. There is a huge difference in the size of the predicted price increase in the two graphs. In the top graph, the oil price increases only a little. If the elasticity is very high, then only a small increase in the

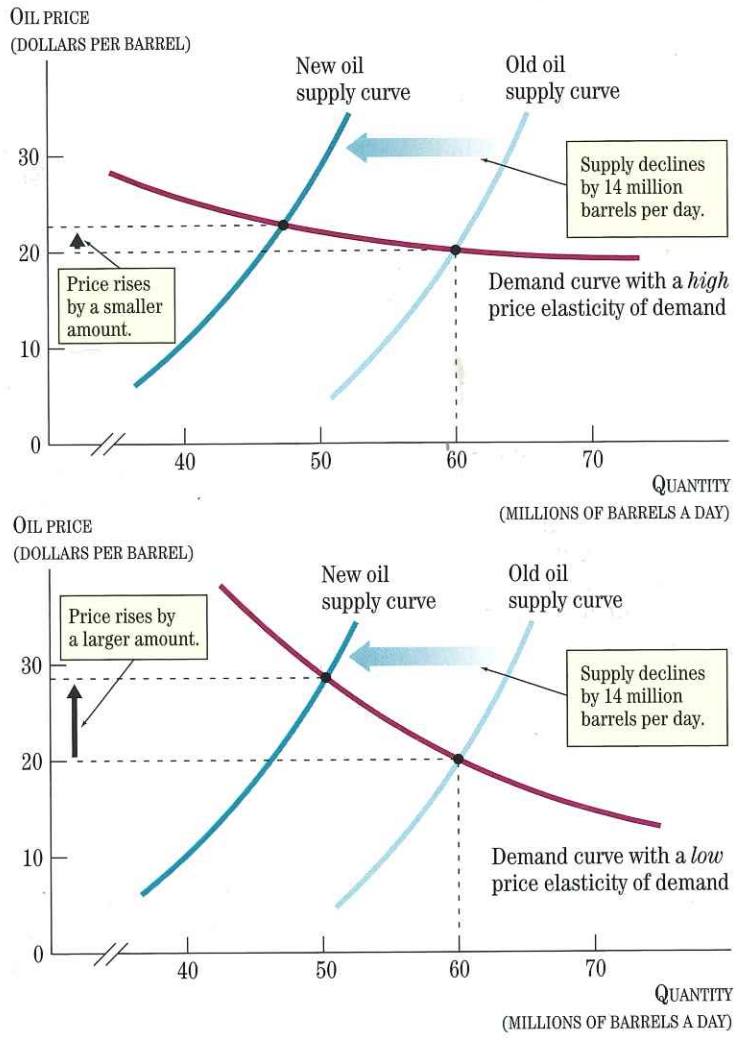


Figure 4.2
The Importance of the Size of the Price Elasticity of Demand
The impact on the oil price of a reduction in oil supply is shown for two different demand curves. The reduction in supply is the same for both graphs. When the price elasticity of demand is high (top graph), there is only a small increase in the price. When the price elasticity of demand is low (bottom graph), the price rises by much more.

price is enough to get people to reduce their use of oil and thereby bring the quantity demanded down to a lower quantity supplied. On the other hand, in the bottom diagram, the price rises by much more. Here the elasticity is very low, and so a large increase in price is needed to get people to reduce their use of oil and bring the quantity demanded down to the quantity supplied.

Thus, in order to determine how much the price will rise in response to a shift in oil supply, we need to know how sensitive the quantity demanded is to the price, or the size of the elasticity of demand.

REVIEW

- The price elasticity of demand is a number that tells us how sensitive the quantity demanded is to the price. It is defined as the percentage change in the quantity demanded divided by the percentage change in the price.

Predicting the Size of a Price Increase

Economists used a numerical value of elasticity to predict the size of the oil price rise caused by the Iraqi invasion of Kuwait in 1990. Here are the steps they took:

- First, they determined—after looking at historical studies of oil prices and quantities—that the price elasticity of the demand for oil was .1. In other words, $e_d = .1$.
- Second, they calculated—after consulting with oil producers—that the invasion of Kuwait would reduce the world oil supply by 7 percent. They assumed that this 7 percent would also be the percentage decline in the quantity of oil demanded because other sources of oil could not increase in a short period of time. In other words, $\Delta Q_d/Q_d = .07$, or 7 percent.
- Third, they plugged these numbers into the formula for elasticity to calculate that the oil price would rise by 70 percent. Here is the exact calculation behind this step: Rearrange the definition of elasticity, $e_d = (\Delta Q_d/Q_d)/(\Delta P/P)$, to put the percentage change in the price on the left. That is, $\Delta P/P = (\Delta Q_d/Q_d)/e_d$. Now plug in $\Delta Q_d/Q_d = .07$ and $e_d = .1$ to get $.07/(.1) = .70$, or 70 percent.

The 70 percent price rise predicted might seem large. In fact, the actual rise in the price of oil in 1990 *was* large, even larger than 70 percent: The price of oil rose from $17 per barrel in July 1990 to $36 in October 1990, or about 112 percent. (The larger-than-predicted price increase may have been due to worries that Iraq would also invade Saudi Arabia and reduce the oil supply even further.)

This type of calculation—showing that a huge oil price increase could be caused by the 7 percent reduction in oil supply—was a factor in the decision by the United States and its allies to send troops to the Middle East to halt the Iraqi invasion of Saudi Arabia and to eventually force Iraq out of Kuwait.

Could you use the same type of reasoning to determine how much the price of oil would *fall* if oil producers *increased* supply? Suppose the increase was 4 percent.

Working with Demand Elasticities

Having demonstrated the practical importance of elasticity, let us examine the concept in more detail and show how to use it. Some symbols will be helpful.

If we let the symbol e_d represent the price elasticity of demand, then we can write the definition as

$$e_d = \frac{\Delta Q_d}{Q_d} \div \frac{\Delta P}{P} = \frac{\Delta Q_d/Q_d}{\Delta P/P}$$

where Q_d is the quantity demanded, P is the price, and Δ means "change in." In other words, the elasticity of demand equals the "percentage change in the quantity demanded" divided by the "percentage change in the price." Observe that to compute the percentage change in the numerator and the denominator, we need to divide the change in the variable (ΔP or ΔQ_d) by the variable (P or Q_d).

Because the quantity demanded is negatively related to the price along a demand curve, the elasticity of demand is a negative number: When $\Delta P/P$ is positive, $\Delta Q_d/Q_d$ is negative. But when economists write or talk about elasticity, they usually ignore the negative sign and report the absolute value of the number. Because the demand curve always slopes downward, this nearly universal convention need not cause any confusion, as long as you remember it.

It is easy to do back-of-the-envelope computations of price elasticity of demand. When the price of Australian wine fell 8.5 percent in 2004, the quantity sold increased 11 percent. The price elasticity of demand for Australian wine is $11/8.5 = 1.3$. At the University of Michigan, when student season ticket prices for hockey games rose 48

percent in 1998, student purchases of season tickets fell 41.1 percent. The price elasticity of demand for season hockey tickets is 41.1/48 = 0.86. Notice that measured in percentage changes, the demand for Australian wine is responsive to changes in the price, and the demand for hockey season tickets is not very responsive to changes in the price. Is this what you'd expect?

The Advantage of a Unit-Free Measure

unit-free measure: a measure that does not depend on a unit of measurement.

An attractive feature of the price elasticity of demand is that it does not depend on the units of measurement of the quantity demanded—whether barrels of oil or pounds of peanuts. It is a **unit-free measure** because it uses *percentage changes* in price and quantity demanded. Thus, it provides a way to compare the price sensitivity of the demand for many different goods. It even allows us to compare the price sensitivity of less expensive goods—like rice—with that of more expensive goods—like steak.

For example, suppose that when the price of rice rises from 50 cents to 60 cents per pound, the quantity demanded falls from 20 tons to 19 tons: That is a decline of 1 ton for a 10 *cent* price increase.

In contrast, suppose that when the price of steak rises by $1, from $5 to $6 per pound, the quantity demanded falls by 1 ton, from 20 tons to 19 tons of steak. That would be a decline of 1 ton for a 1 *dollar* price increase.

Using these numbers, the price sensitivity of the demand for steak and the demand for rice might appear to be very different: 10 cents to get a ton of reduced purchases versus $1 to get a ton of reduced purchases. Yet the elasticities are the same. The percentage change in price is 20 percent in each case ($1/$5 = $.10/$.50 = .20, or 20 percent), and the percentage change in quantity is 5 percent in each case: 1 ton of rice/20 tons of rice = 1 ton of steak/20 tons of steak = .05, or 5 percent. Hence, the elasticity is 5/20 = 1/4 in both cases.

Elasticity allows us to compare the price sensitivity of different goods by looking at ratios of percentage changes regardless of the units for measuring either price or quantity. With millions of different goods and hundreds of different units of measurement, this is indeed a major advantage.

Elasticity versus Slope

The *elasticity of the demand curve* is not the same as the *slope of the demand curve.* The slope of the demand curve is defined as the change in price divided by the change in quantity demanded: $\Delta P/\Delta Q$. The slope is not unit-free; it depends on how the price and quantity are measured. Thus, it is not a good measure of price sensitivity when we compare different goods.

To illustrate the difference between slope and elasticity, we show in Figure 4.3 a demand curve for rice and a demand curve for steak. The two demand curves have different slopes because the prices are so different. In fact, the slope of the steak demand curve is 10 times greater than the slope of the rice demand curve. Yet the elasticity is the same for the change from *A* to *B* for both demand curves.

After looking at Figure 4.1, you might be tempted to say that demand curves that are very steep have a low elasticity, and demand curves that are very flat have a high elasticity. However, because the slope and the elasticity are such different concepts, we must be careful not to simply look at a steep curve and say that it has a high elasticity. Moreover, as we show shortly, a curve with a constant slope has a different elasticity at every point on the curve!

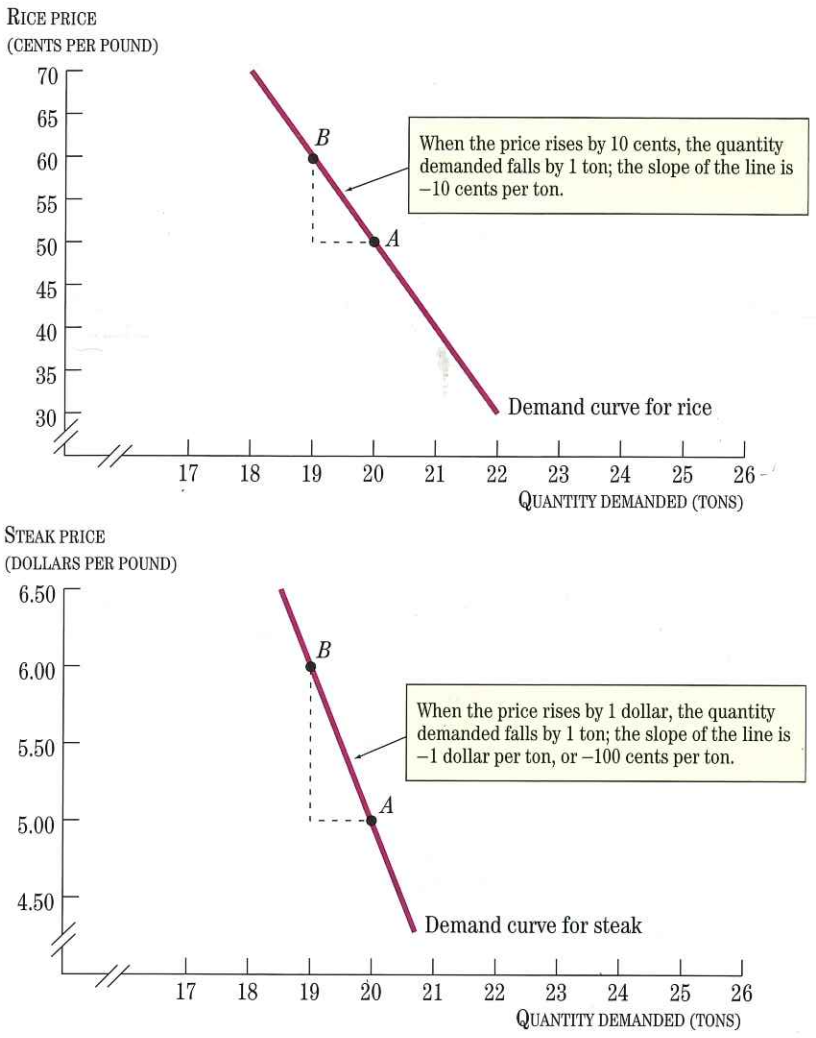


Figure 4.3
Different Slopes and Same Elasticities
The slope of the steak demand curve in the bottom graph is greater than the slope of the rice demand curve in the top graph. The price elasticity of demand for rice and steak from point *A* to point *B* is the same, however. From point *A* to point *B*, the price rises by 20 percent and the quantity demanded decreases by 5 percent. Thus, the elasticity is 1/4 for both rice and steak at these points.

Talking About Elasticities

Economists classify demand curves by the size of the price elasticities of demand, and they have developed a very precise terminology for doing so.

elastic demand: demand for which the price elasticity is greater than 1.

inelastic demand: demand for which the price elasticity is less than 1.

■ **Elastic versus Inelastic Demand.** Goods for which the price elasticity is greater than 1 have an **elastic demand.** For example, the quantity of foreign travel demanded decreases by more than 1 percent when the price rises by 1 percent because many people tend to travel at home rather than abroad when the price of foreign travel rises.

Goods for which the price elasticity of demand is less than 1 have an **inelastic demand.** For example, the quantity of eggs demanded decreases by less

than 1 percent when the price of eggs rises by 1 percent because many people do not want to substitute other things for eggs at breakfast.

perfectly inelastic demand: demand for which the price elasticity is zero, indicating no response to a change in price and therefore a vertical demand curve.

perfectly elastic demand: demand for which the price elasticity is infinite, indicating an infinite response to a change in the price and therefore a horizontal demand curve.

■ **Perfectly Elastic versus Perfectly Inelastic Demand.** A demand curve that is vertical is called **perfectly inelastic.** Figure 4.4 shows a perfectly inelastic demand curve. The elasticity is zero because when the price changes, the quantity demanded does not change at all. No matter what the price, the same quantity is demanded. People who need insulin have a perfectly inelastic demand for insulin. As long as there are no substitutes for insulin, they will pay whatever they have to in order to get the insulin.

A demand curve that is horizontal is called **perfectly elastic.** Figure 4.4 also shows a perfectly elastic demand curve. The elasticity is infinite. The perfectly flat demand curve is sometimes hard to imagine because it entails infinitely large movements of quantity for tiny changes in price. In order to better visualize this case, you can imagine that the curve is tilted ever so slightly. The infinity case is extreme and is used to approximate demand curves with very high elasticities.

Table 4.1 summarizes the terminology about elasticities.

Calculating the Elasticity with a Midpoint Formula

To calculate the elasticity, we need to find the percentage change in the quantity demanded and divide it by the percentage change in the price. As we have already illustrated with examples, to get the percentage change in the price or quantity, we need to divide the change in price (ΔP) by the price (P) or the change in quantity demanded (ΔQ_d) by the quantity demanded (Q_d). But when price and quantity demanded change, there is a question about what to use for P and Q_d. Should we use the old price and old quantity demanded before the change, or should we use the new price and new quantity demanded after the change?

The most common convention economists use is a compromise between these two alternatives. They take the *average,* or the *midpoint,* of the old and new quantities demanded and the old and new prices. That is, they compute the elasticity using the following formula, called the *midpoint formula:*

$$\text{Price elasticity of demand} = \frac{\text{change in quantity}}{\text{average of old and new quantities}} \div \frac{\text{change in price}}{\text{average of old and new prices}}$$

For example, if we use the midpoint formula to calculate the price elasticity of demand for oil when the price changes from $20 to

Figure 4.4
Perfectly Elastic and Perfectly Inelastic Demand
A perfectly inelastic demand curve is a vertical line at a certain quantity; the quantity demanded is completely insensitive to the price: Whatever happens to the price, the quantity demanded does not change. A perfectly elastic demand curve is a flat line at a certain price; an increase in price reduces the quantity demanded to zero; a small decrease in price raises the quantity demanded by a huge (literally infinite) amount.

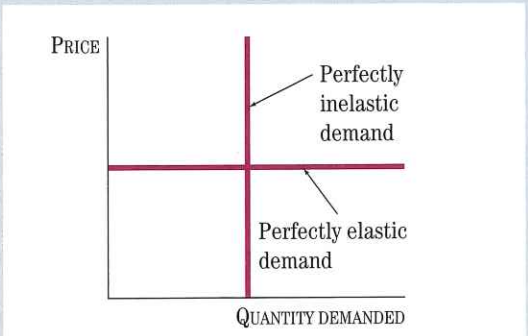


Table 4.1
Terminology for Price Elasticity of Demand

Term	Value of Price Elasticity of Demand (e_d)
Perfectly inelastic	0 (vertical demand curve)
Inelastic	Less than 1
Elastic	Greater than 1
Perfectly elastic	Infinity (horizontal demand curve)

Creating and Advancing Elasticity

Alfred Marshall created the powerful unit-free concept called elasticity, for which other economists are eternally grateful. The great economist John Maynard Keynes said, "I do not think Marshall did economists any greater service." Marshall defined elasticity in a remarkable book, *Principles of Economics*, published exactly 100 years before the economists used elasticity to predict the oil price increase when Iraq invaded Kuwait. Here is how Marshall put it:

> The elasticity of demand in a market is great or small according as the amount demanded increases much or little for a given fall in price, and diminishes much or little for a given rise in price.

Marshall's *Principles of Economics* had an enormous impact on economic thinking and thus on economic policy. More than anyone else, Marshall showed how to use the supply and demand model to analyze real-world problems. "Economics in Action" could have been his motto. Many economists credit Marshall with developing the *style* of economic reasoning that is routinely followed in applied economics work today. Marshall taught economists how to think like economists.

Marshall was born south of London to a family that was not particularly well-off; his father was a dominating figure and forbade Alfred to study mathematics, which Marshall loved but his father did not understand. Rather than go to Oxford, where his father wanted him to go, Marshall went to Cambridge.

Marshall did not decide to become an economist until after college. The reason for his choice? He said he saw economics as a way to help the poor and unfortunate in society. Marshall was the greatest economist of the neoclassical school of the late nineteenth and early twentieth centuries. The neoclassical school maintained most of the principles of the classical school founded by Adam Smith, but improved on them with the supply and demand model and its underlying foundation.

Marshall's simple, yet broad, definition of economics has itself become a classic. To Marshall,

> Political Economy or Economics is a study of mankind in the ordinary business of life; it examines that part of individual and social action which is most closely connected with the attainment and with the use of the material requisites of wellbeing.

Based on your study of economics so far, do you think Marshall's definition is a good one?

$22 and the quantity demanded changes from 60 million to 48 million barrels a day, we get

$$\frac{60 - 48}{(60 + 48)/2} \div \frac{\$20 - \$22}{(\$20 + \$22)/2} = .2222 \div (-.0952) = -2.33$$

That is, the price elasticity of demand is 2.33 using the midpoint formula. In this case, this is close to the value of 2 we obtained by using the old price ($20) rather than the average, $21 = ($20 + $22)/2, and the old quantity demanded (60) rather than the average quantity, 54 = (60 + 48)/2.

Revenue and the Price Elasticity of Demand

When people purchase 60 million barrels of oil at $20 a barrel, they must pay a total of $1,200 million ($20 × 60 million). This is a payment to the oil producers and is the producers' revenue. In general, revenue is the price (P) times the quantity (Q), or $P \times Q$.

The Two Effects of Price on Revenue. Because revenue is defined as price times quantity, an increase or a decrease in the price has an effect on revenue. However, there are two opposite effects. Consider an increase in the price. An increase in the price raises the payment per unit but also reduces the number of units. In other words, when the price increases, people pay more for each item, and

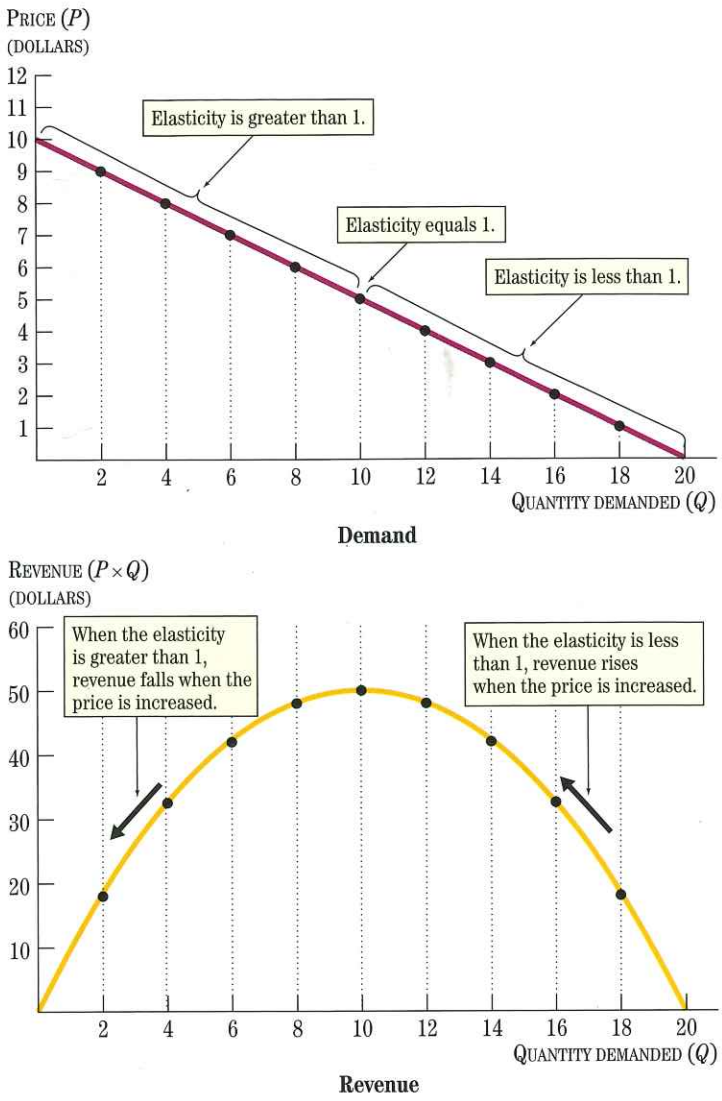


Figure 4.5
Revenue and Elasticity of a Straight-Line Demand Curve
Along the straight-line demand curve at the top, the price elasticity ranges from above 1 (to the left) to below 1 (to the right). When the price elasticity is greater than 1, an increase in the price will reduce revenue, as shown in the lower panel.

this increases revenue; but they buy fewer items, and this decrease in the quantity demanded reduces revenue. The elasticity determines which of these two opposite effects dominates.

Figure 4.5 illustrates how the price elasticity determines the effect of a price change on revenue for the special case of a demand curve that is a straight line. Look first at the demand curve shown in the top panel of Figure 4.5. Observe that the elasticity is greater than 1 on the left part of the demand curve and less than 1 on the right part of the demand curve.

Now look at the lower panel in Figure 4.5; it shows how revenue is related to price in the case of this same straight-line demand curve. Observe how revenue changes as the price increases. When the price elasticity is greater than 1—in the region on the left—an increase in the price will lower revenue. For example, an increase in the price

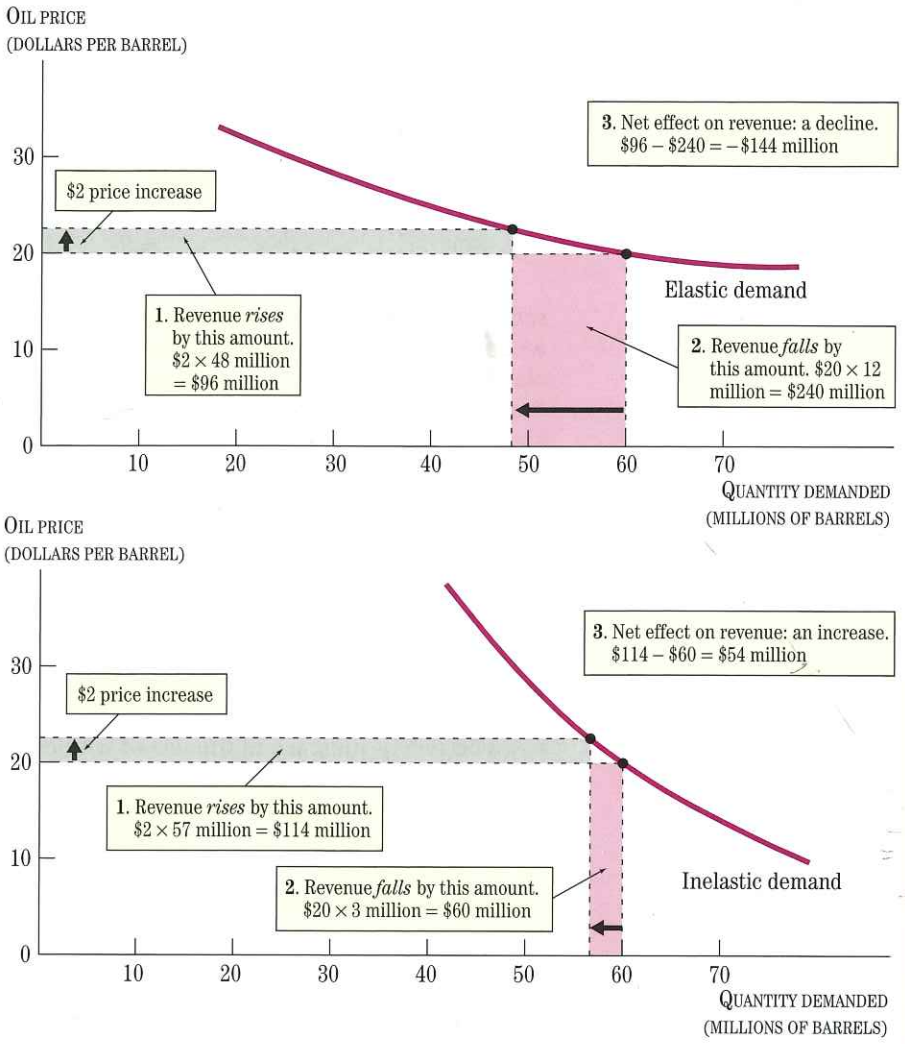


Figure 4.6
Effects of an Increase in the Price of Oil on Revenue
These graphs are replicas of the demand curves for oil shown in Figure 4.1, with the scale changed to show the change in revenue when the price of oil is increased. An increase in the price has two effects on revenue, as shown by the gray- and pink-shaded rectangles. The increase in revenue (gray rectangle) is due to the higher price. The decrease in revenue (pink rectangle) is due to the decline in the quantity demanded as the price is increased. In the top graph, where elasticity is greater than 1, the net effect is a decline in revenue; in the bottom graph, where elasticity is less than 1, the net effect is an increase in revenue.

from $8 per unit to $9 per unit reduces the quantity demanded from 4 units to 2 units and revenue falls from $32 to $18; the gain in revenue from the higher price per unit is offset by the loss in revenue from the decline in the number of units sold. However, when the price elasticity is less than 1—in the region on the right—an increase in the price will raise revenue. For example, an increase in the price from $1 to $2 per unit reduces the quantity demanded from 18 units to 16 units and revenue rises from $18 to $32; in this case, the loss in revenue from the decline in the number of units sold is not large enough to offset the gain in revenue from the higher price per unit.

Another illustration of the relationship between elasticity and revenue changes is found in Figure 4.6, which is a replica of Figure 4.1 with the scales changed to better illustrate the effects on revenue. We now know that the elasticity in the top graph is greater than 1 and the elasticity in the bottom graph is less than 1. We can

Table 4.2
Revenue and the Price Elasticity of Demand

Elasticity Is	Effect of a Price Increase on Revenue	Effect of a Price Decrease on Revenue
Less than 1 (< 1)	Revenue increases	Revenue decreases
Equal to 1 (= 1)	No change in revenue	No change in revenue
Greater than 1 (> 1)	Revenue decreases	Revenue increases

see that the same price increase leads to a large decline in the quantity demanded when the elasticity is large and to a small decline in the quantity demanded when the elasticity is small. The two effects of a price increase are illustrated in the graphs. Revenue falls when the price is increased in the top graph and rises when the price is increased in the bottom graph.

Table 4.2 summarizes the relationship between revenue and the price elasticity of demand. Observe that an increase in price will raise revenue if the elasticity is less than 1 and will lower revenue if the elasticity is greater than 1.

Businesses need to know the price elasticity of demand for their products. The most recent U.S. recession led to a decrease in demand for business air travel. In 2003, United Airlines announced 40 percent cuts in some one-way business fares. Would more businesspeople decide to fly with lower fares? If so, would the increase in customers lead to an increase or a decrease in United Airlines' revenue? The answer depends on the price elasticity of demand for business air travel.

The recent increase in the use of wireless phones and prepaid phone cards has resulted in a decrease in the demand for pay phones. In 2001, because of this decrease in demand, SBC Communications Inc. increased the price of a pay phone call. Could this increase in price lead to an increase in revenue? The answer depends on the price elasticity of demand for pay phone calls. Notice that an increase in price could increase revenue, but a decrease in price could also increase revenue, depending on the price elasticity of demand.

If demand for business air travel is price elastic and demand for pay phones is price inelastic, then both United Airlines and SBC changed prices to increase revenue—United Airlines cutting business fares to increase revenue with price-elastic demand, and SBC increasing the price of using a pay phone call to increase revenue with price-inelastic demand. The next section discusses the determinants of price elasticity of demand. You should judge as you read this chapter whether demand for business air travel and demand for pay phone calls are likely to be price elastic or price inelastic.

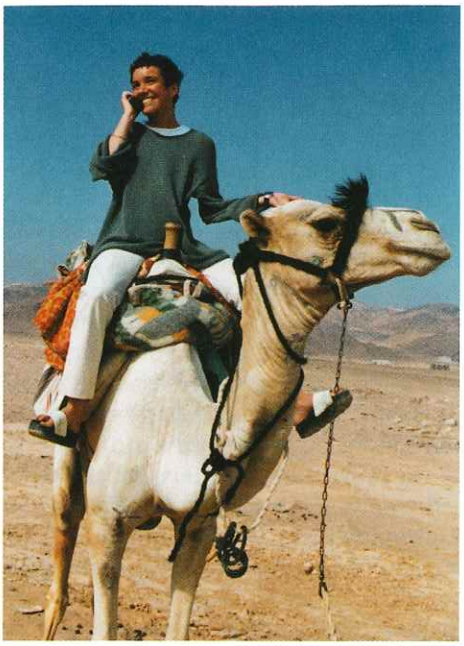
Table 4.3
Estimated Price Elasticities of Demand

Type of Good or Service	Price Elasticity
Jewelry	2.6
Eggs	0.1
Telephone (first line)	0.1
Telephone (second line)	0.4
Foreign travel	1.2
Cigarettes (18–24)	0.6
Cigarettes (25–39)	0.4
Cigarettes (40–older)	0.1
Gasoline (short run)	0.2
Gasoline (long run)	0.7

What Determines the Size of the Price Elasticity of Demand?

Table 4.3 shows price elasticities of demand for several different goods and services. The price elasticity for jewelry, for example, is 2.6. This means that for each percentage increase in the price of jewelry, the quantity demanded will fall by 2.6 percent. Compared with other elasticities, this is large. On the other hand, the price elasticity of eggs is very small. For each percentage increase in the price of eggs, the quantity of eggs demanded falls by only 0.1 percent.

Why do these elasticities differ in size? Several factors determine a good's elasticity.



Price Elasticity of Cell Phone Service in the Desert
There's probably not much of a substitute available for long-distance communication in the desert, which would make the price elasticity of the cell phone used here quite low. Do you think the caller in this picture would be equally insensitive to an increase of $0.10 per minute in the price of her calls if she were seated in her apartment in Chicago?

■ **The Degree of Substitutability.** A key factor is whether there are good substitutes for the item in question. Can people easily find a substitute when the price goes up? If the answer is yes, then the price elasticity will be high. Foreign travel has a high elasticity because there is a reasonably good substitute: domestic travel.

On the other hand, the low price elasticity for eggs can be explained by the lack of good substitutes. As many fans of eggs know, these items are unique; synthetic eggs are not good substitutes. Hence, the price elasticity of eggs is small. People will continue to buy them even if the price rises a lot.

The degree of substitutability depends in part on whether a good is a necessity or a luxury. There are no good substitutes for a refrigerator if you want to easily preserve food for more than a few hours. However, a fancy refrigerator with an exterior that blends in with the rest of your kitchen is more of a luxury and is likely to have a higher price elasticity.

■ **Big-Ticket versus Little-Ticket Items.** If a good represents a large fraction of people's income, then the price elasticity will be high. If the price of foreign travel doubles, many people will not be able to afford to travel abroad. On the other hand, if the good represents a small fraction of income, the elasticity will be low. For example, if the price of eggs doubles, most people will still be able to afford to buy as many eggs as before the price rise.

■ **Temporary versus Permanent Price Changes.** If a change in price is known to be temporary, the price elasticity of demand will tend to be high because many people can easily shift their purchases either later or earlier. For example, suppose a sewing machine store announces a discount price to last only one day. Then people will shift their purchase of the sewing machine they were thinking about buying to the sale day.

On the other hand, if the price cut is permanent, the price elasticity will be smaller. People who expect the price decrease to be permanent will not find it advantageous to buy sooner rather than later.

■ **Differences in Preferences.** Various groups of consumers may have different levels of elasticity. For example, young cigarette smokers, whose habit of smoking may not be entrenched, are more sensitive to changes in prices than older smokers. Table 4.3 shows how the price elasticity of demand for cigarettes for young adults between 18 and 24 years old is much higher than the very low price elasticity for people older than 40.

■ **Long-Run versus Short-Run Elasticity.** Frequently the price elasticity of demand is low immediately after a price change but then increases after a period of time has passed. In order to analyze these changes, economists distinguish between the *short run* and the *long run*. The short run is simply a period of time before people have made all their adjustments or changed their habits; the long run is a period of time long enough for people to make such adjustments or change their habits.

Many personal adjustments to a change in prices take a long time. For example, when the price of gas increases, people can reduce the quantity demanded in the short run only by driving less and using other forms of transportation more, or by reducing the heating in their homes. This may be inconvenient or impossible. In the long run, however, when it comes time to buy a new car or a new heating system, they can buy a more fuel-efficient one, or one that uses an alternative energy source. Thus, the quantity of gas demanded falls by larger amounts in the long run than in the short run (Table 4.3).

What is the Price Elasticity of Demand for Star Wars Movies?
For these fans, the prospect of attending the May 2005 opening of Star Wars: Episode III—Revenge of the Sith *is price inelastic (i.e., they will pay almost anything for a ticket to this opening); but for moviegoers in general, the answer to that question may say a lot about the future of the movie industry. If purchases of movie tickets fall off as the price of a general admission ticket goes up, and at the same time revenue from ticket sales increases, then demand for movie tickets is price inelastic.*

Habits that are difficult to break also cause differences between short-run and long-run elasticity. Even a large increase in the price of tobacco may have a small effect on the quantity purchased because people cannot break the smoking habit quickly. But after a period of time, the high price of cigarettes may encourage them to break the habit, while discouraging potential new users. Thus, the long-run elasticity for tobacco is higher than the short-run elasticity.

Here are a few examples to test your understanding of the determinants of the price elasticity of demand. The movie industry reported that its summer 2004 revenue was 3 percent higher than the previous year. A closer analysis reveals that ticket sales were down 1 percent. How could ticket revenue increase at the same time that the number of tickets sold decreased? The ticket price must have increased. Demand for movies must also be price inelastic, so that the reduction in ticket sales was more than offset by the increase in the price of the movie tickets. Does this make sense for the movie industry? It is plausible that some people feel they *must see* the newest release and that the cost of the movie is a little-ticket item for many people who go to the movies. This would make the price elasticity of demand low and demand plausibly price inelastic.

General Motors Corporation reported that in 2002 its revenue rose by 5.4 percent. At the same time, General Motors offered large discounts to customers purchasing cars. How could revenue increase while the price of cars was going down? It must be that more cars were sold at the lower price and that demand for these cars is price elastic. The reduction in price was therefore offset by the increase in cars sold, and revenue increased. Does this make sense for General Motors cars? It is plausible that customers feel there are close substitutes and that this is a big-ticket purchase for many customers. This would make the price elasticity of demand high and demand plausibly price elastic.

Income Elasticity and Cross-Price Elasticity of Demand

Recall that the price elasticity of demand refers to movements along the demand curve. We emphasized in Chapter 3 the difference between a shift in the demand curve and a movement along the demand curve. A *shift* in the demand curve occurs when there is a change in the quantity people are willing to buy due to a change in anything except the price—for example, a change in income.

income elasticity of demand: the percentage change in quantity demanded of a good divided by the percentage change in income.

The concept of elasticity can be applied to changes in the quantity consumers are willing to buy caused by changes in income. This elasticity must be distinguished from the price elasticity of demand. The **income elasticity of demand** is the percentage change in the quantity of a good demanded at any given price divided by a percentage change in income. That is,

$$\text{Income elasticity of demand} = \frac{\text{percentage change in quantity demanded}}{\text{percentage change in income}}$$

Table 4.4
Estimated Income Elasticities of Demand

Type of Good or Service	Income Elasticity
Food	0.58
Clothing/footwear	0.88
Transport	1.18
Medical Care	1.35
Recreation	1.42

For example, if incomes rise by 10 percent and, as a result, people purchase 15 percent more health care at a given price, the income elasticity of health care is 1.5. Table 4.4 lists income elasticities of demand for several different goods and services.

As discussed in Chapter 3, the demand for most goods increases when people's incomes increase. If you have more income, your demand for movies will probably increase at each price. Recall that a normal good is a good or service whose demand increases as income increases. But not every good is a normal good; if the demand

for a good declines when income increases, the good is called an inferior good. The income elasticity of demand for an inferior good is negative and is reported as a negative number by economists.

cross-price elasticity of demand: the percentage change in the quantity demanded of one good divided by the percentage change in the price of another good.

Another type of elasticity relating to shifts in the demand curve is the **cross-price elasticity of demand,** which is defined as the percentage change in the quantity demanded divided by the percentage change in the price of another good. For example, an increase in the price of Rollerblades would *increase* the quantity demanded of bicycles at every price as people shifted away from Rollerblading to bicycle riding. Rollerblades are a substitute for bicycles. A cross-price elasticity can also go in the other direction. An increase in the price of bicycle helmets may *reduce* the demand for bicycles. Bicycle helmets and bicycles are complements. For a complement, the cross-price elasticity of demand is negative.

REVIEW

- The price elasticity of demand is a unit-free number that is different from the slope of the demand curve.
- The elasticity helps determine how large a price increase will occur as a result of a shift in supply, and how much revenue will change when the price rises.
- Horizontal demand curves have infinite price elasticity. Vertical demand curves have zero price elasticity. Most products have a price elasticity between these two extremes.
- The size of the price elasticity of demand depends on the availability of substitutes for the item, whether the item represents a large fraction of income, and whether the price change is temporary or permanent.
- Whereas the price elasticity of demand refers to movements along the demand curve, the income elasticity of demand refers to shifts in the demand curve caused by changes in income. Most goods are normal and have a positive income elasticity of demand. Inferior goods have a negative income elasticity of demand. The cross-price elasticity of demand also relates to shifts in the demand curve, in this case, a change in the price of a complement or substitute good.

Elasticity of Supply

Knowing how sensitive the quantity supplied is to a change in price is just as important as knowing how sensitive the quantity demanded is. The price elasticity of supply measures this sensitivity. "Price elasticity of supply" is sometimes shortened to "supply elasticity" or "elasticity of supply." Supply describes the behavior of firms that produce goods. A high price elasticity of supply means that firms raise their production by a large amount if the price increases. A low price elasticity of supply means that firms raise their production only a little if the price increases.

price elasticity of supply: the percentage change in quantity supplied divided by the percentage change in price.

The **price elasticity of supply** is defined as the percentage change in the quantity supplied divided by the percentage change in the price. That is,

$$\text{Price elasticity of supply} = \frac{\text{percentage change in quantity supplied}}{\text{percentage change in the price}}$$

The price elasticity of supply refers to a particular supply curve, such as the supply curve for gasoline or video games. All other things that affect supply are held constant when we compute the price elasticity of supply. For example, suppose the price elasticity of supply for video games is .5. Then, if the price of video games rises by 10 percent, the quantity of video games supplied will increase by 5 percent (.5 × 10).

Working with Supply Elasticities

All the attractive features of the price elasticity of demand also apply to the price elasticity of supply. To see this, let us first take a look at the definition of the price elasticity of supply using symbols. If we let the symbol e_s be the price elasticity of supply, then it can be written as

$$e_s = \frac{\Delta Q_s}{Q_s} \div \frac{\Delta P}{P} = \frac{\Delta Q_s / Q_s}{\Delta P / P}$$

where Q_s is the quantity supplied and P is the price. In other words, the price elasticity of supply is the percentage change in the quantity supplied divided by the percentage change in price. Observe the similarity of this expression to the analogous expression for the price elasticity of demand on page 89: The only difference is the use of quantity supplied (Q_s) rather than quantity demanded (Q_d). This means that the concepts and terminology for supply elasticity are very similar to those for demand elasticity. For example, if you go to Table 4.1 and replace "Demand" with "Supply," you have the terminology of price elasticity of supply. Moreover, like the price elasticity of demand, the price elasticity of supply is a unit-free measure, and the elasticity of supply and the slope of the supply curve are not the same thing.

Because of this similarity, our discussion of supply elasticity can be very brief. It is useful to consider the extreme cases of perfectly elastic supply and perfectly

Perfectly Inelastic Supply
The paintings of Leonardo Da Vinci provide an example of a good with a perfectly inelastic supply. The supply curve is vertical because no matter how high the price, no more Mona Lisas can be produced. However, what about the demand to see the Mona Lisa? Is it perfectly inelastic? Will raising the price of admission charged by the Louvre Museum in Paris reduce the number of people coming to see the painting?

inelastic supply, and then to go through an example illustrating the importance of knowing the size of the price elasticity of supply.

perfectly elastic supply: supply for which the price elasticity is infinite, indicating an infinite response of quantity supplied to a change in price and thereby a horizontal supply curve.

perfectly inelastic supply: supply for which the price elasticity is zero, indicating no response of quantity supplied to a change in price and thereby a vertical supply curve.

■ **Perfectly Elastic and Perfectly Inelastic Supply.** As in the case of demand, there can be **perfectly elastic supply** or **perfectly inelastic supply,** as shown in Figure 4.7. The vertical supply curve is perfectly inelastic; it has zero elasticity. Such supply curves are not unusual. For example, there is only one *Mona Lisa.* A higher price cannot bring about a higher quantity supplied, not even one more *Mona Lisa.* But the supply of most goods is not vertical. Higher prices will encourage coffee producers to use more fertilizer, hire more workers, and eventually plant more coffee trees. Thus the quantity supplied increases when the price rises.

The horizontal supply curve is perfectly elastic. In this case, the price does not change at all. It is the same regardless of the quantity supplied. It is easier to understand the horizontal supply curve if you view it as an approximation to a supply curve that is *nearly* horizontal, one with a very high elasticity. Then only a small increase in price brings forth a huge increase in the quantity supplied by firms.

■ **Why the Size of the Price Elasticity of Supply Is Important.** Now let us look at the importance of knowing the size of the supply elasticity even if it is not at one of these two extremes. Figure 4.8 shows two different supply curves for coffee. The horizontal axis shows the quantity of coffee supplied around the world in billions of pounds; the vertical axis shows the price in dollars per pound of coffee. For the supply curve in the top graph, the quantity supplied is very sensitive to the price;

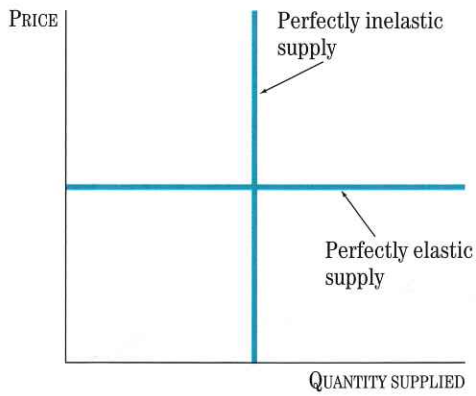


Figure 4.7
Perfectly Elastic and Perfectly Inelastic Supply
When the quantity supplied is completely unresponsive to the price, the supply curve is vertical and the price elasticity of supply is zero; this case is called perfectly inelastic supply. When the quantity supplied responds by large amounts to a price change, the supply curve is horizontal; economists then say supply is perfectly elastic.

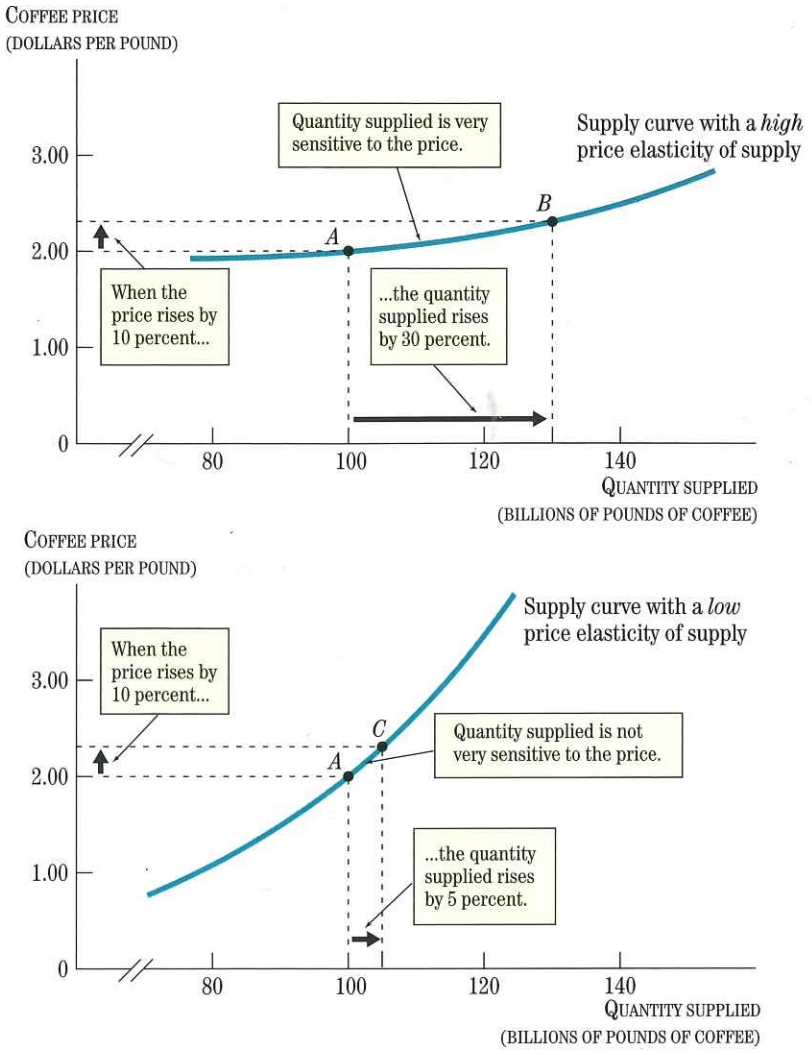


Figure 4.8
Comparing Different Sizes of the Price Elasticities of Supply
In the top graph, the quantity supplied is much more sensitive to price than in the bottom graph. The price elasticity of supply is greater between points *A* and *B* at the top than between points *A* and *C* at the bottom.

the price elasticity of supply is high. For the supply curve in the bottom graph, the price elasticity of supply is much lower.

The price elasticity of supply is important for finding out the response of price to shifts in demand. This is shown in Figure 4.9, where the demand for coffee declines, perhaps because of concerns about the effect of the caffeine in coffee or because of a decrease in the price of caffeine-free substitutes for coffee. In any case, if the price elasticity of supply is high, as in the top graph, the price does not change as much as when the price elasticity of supply is low, as in the bottom graph. With a high price elasticity, a small change in price is enough to get firms to bring the quantity supplied down to the lower quantity demanded.

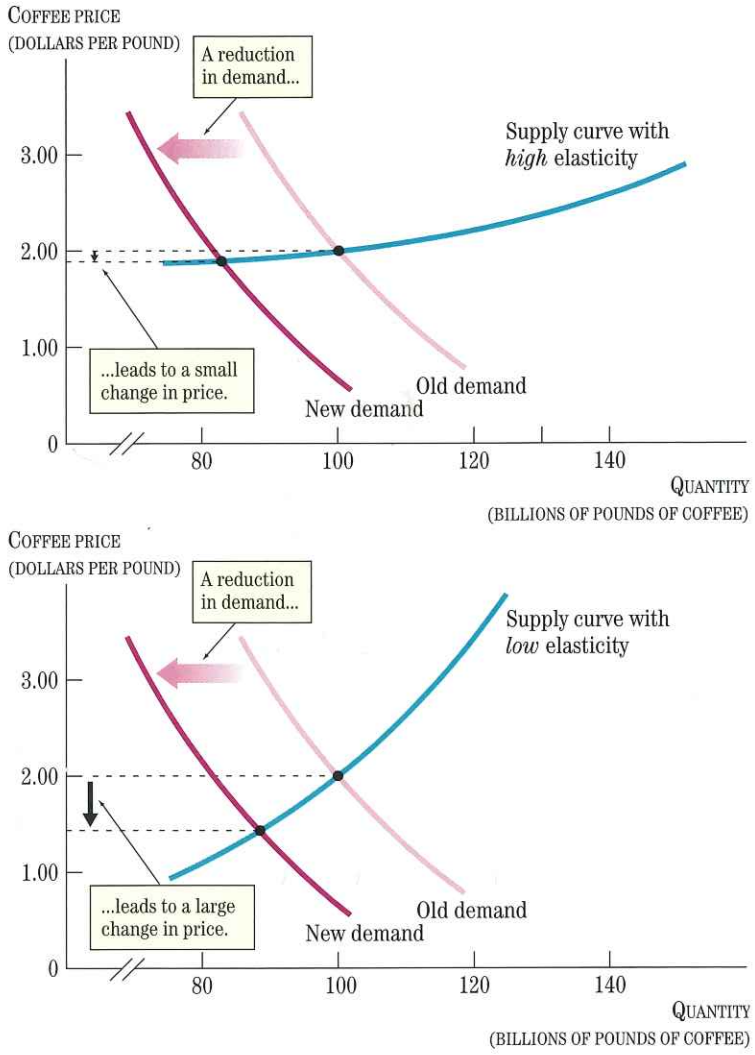


Figure 4.9
Importance of Knowing the Size of the Price Elasticity of Supply
When demand changes, the price will also change. If the price elasticity of supply is high, there will be a small change in price. If the price elasticity of supply is low, there will be a large change in price.

REVIEW

- The price elasticity of supply is a number that tells us how sensitive the quantity supplied is to the price. It is defined as the percentage change in the quantity supplied divided by the percentage change in the price.
- The attractive features of the price elasticity of demand are also true for the price elasticity of supply. Its size does not depend on the units of measurement of either price or quantity.
- The price elasticity of supply is useful for determining how much prices will change when there is a change in demand.

Conclusion

In this chapter, we have seen that *how much* the equilibrium price and quantity change in response to changes in supply or demand depends on the size of the elasticity of the supply and demand curves. Armed with these elasticities, we can predict what will happen to prices after a change in the supply of a good. We can also predict what will happen to prices after a change in demand. And we can predict whether revenue will increase or decrease when prices, fees, or taxes are cut or raised.

Distinguishing between movements along and shifts of the curves, knowing the importance of the elasticity, and describing in words what happens when the intersection of the curves changes are all important parts of thinking like an economist.

KEY POINTS

1. Elasticity is a measure of the sensitivity of one economic variable to another. For example, the price elasticity of demand measures how much the quantity demanded changes when the price changes.
2. The price elasticity of demand is the percentage change in the quantity demanded divided by the percentage change in price. It refers to changes in price and quantity demanded along the demand curve, all other things being equal.
3. Elasticity is a unit-free measure.
4. Demand is elastic if the price elasticity of demand is greater than 1 and inelastic if the price elasticity of demand is less than 1.
5. When the elasticity is greater than 1, an increase in the price reduces the quantity demanded by a percentage greater than the percentage increase in the price, thereby reducing revenue. When the elasticity is less than 1, an increase in the price reduces the quantity demanded by a percentage less than the percentage increase in the price, thereby increasing revenue.
6. The elasticity of demand for a good depends on whether the good has close substitutes, whether its value is a large or small fraction of total income, and the time period of the change.
7. If a good has a low price elasticity of demand, then a change in supply will cause a big change in price. Conversely, if a good has a high price elasticity of demand, then a change in supply will cause a small change in price.
8. Whereas the price elasticity of demand refers to movements along the demand curve, the income elasticity of demand refers to shifts in the demand curve caused by changes in income, and the cross-price elasticity of demand refers to shifts in the demand curve caused by changes in the price of other goods. Most goods are normal and have a positive income elasticity of demand. Inferior goods have a negative income elasticity of demand.
9. The price elasticity of supply is defined as the percentage change in the quantity supplied divided by the percentage change in the price.
10. If a good has a low price elasticity of supply, then a change in demand will cause a big change in price. Conversely, if a good has a high price elasticity of supply, then a change in demand will cause a small change in price.

KEY TERMS

price elasticity of demand
unit-free measure
elastic demand
inelastic demand
perfectly inelastic demand
perfectly elastic demand
income elasticity of demand
cross-price elasticity of demand
price elasticity of supply
perfectly elastic supply
perfectly inelastic supply

QUESTIONS FOR REVIEW

1. Why is the price elasticity of demand a unit-free measure of the sensitivity of the quantity demanded to a price change?
2. What factors determine whether the price elasticity of demand is high or low?
3. What is the difference between the price elasticity of demand and the income elasticity of demand?
4. Why is the price elasticity of demand useful for finding the size of the price change that occurs when supply shifts?

5. What is the difference between elastic and inelastic demand?
6. For what values of the price elasticity of demand do increases in the price increase revenue?
7. What is the slope of a perfectly elastic supply curve?
8. Why is the price elasticity of demand lower in the short run than in the long run?
9. What is the income elasticity of demand?
10. If the price elasticity of demand for textbooks is 2 and the price of textbooks increases by 10 percent, by how much does the quantity demanded fall?

PROBLEMS

1. Using the price elasticities of demand in Table 4.3, determine
   a. which items have an elastic demand and which have an inelastic demand
   b. which items have a *perfectly* elastic demand and which have a *perfectly* inelastic demand
   c. the effect on the quantity demanded of a price decrease of 20 percent
   d. the effect on the price of a shift in supply that reduces the quantity demanded by 10 percent (compare the equilibrium after the supply shift with the equilibrium before the supply shift)
   e. for which items revenues will increase when the price increases
2. The price of parking in a downtown parking garage affects the number of people who want to park there. The manager notices that as the hourly parking fee is raised, total revenue falls. Why is this true? Assuming a linear demand curve, what is the elasticity of demand for parking when total revenue begins to fall? To maximize total revenue, at what elasticity will the manager set the price?
3. What does it mean for the supply curve to be perfectly inelastic? How does revenue change as the demand curve shifts?
4. Compare a market in which supply and demand are very (but not perfectly) inelastic to one in which supply and demand are very (but not perfectly) elastic. Suppose the government decides to impose a price floor $1 above the equilibrium price in each of these markets. Compare, diagrammatically, the surpluses that result. In which market is the surplus larger?
5. Calculate the price elasticity of demand for the following goods:
   a. The price of movie theater tickets goes up by 10 percent, causing the quantity demanded to go down by 4 percent.
   b. Computer prices fall by 20 percent, causing the quantity demanded to increase by 15 percent.
   c. The price of apples rises by 5 percent, causing the quantity demanded to fall by 5 percent.
   d. The price of ice cream falls by 6 percent, causing the quantity demanded to rise by 10 percent.
6. Use the following data for a demand curve.

Price	Quantity
11	10
10	20
9	30
8	40
7	50
6	60
5	70
4	80
3	90

   a. Use the midpoint formula to calculate the elasticity between a price of $10 and $11.
   b. Use the midpoint formula to calculate the elasticity between $3 and $4.
   c. Since this is a linear demand curve, why does the elasticity change?
   d. At what point is price times quantity maximized? What is the elasticity at that point?
7. Use the following data to calculate the price elasticity of supply for the price between $7 and $8 and the price between $3 and $4. Use the midpoint formula. How does supply elasticity change as you move up the supply curve?

Price	Quantity Supplied
2	10
3	20
4	30
5	40
6	50
7	60
8	70
9	80

8. Suppose that the demand for cigarettes is perfectly inelastic. If the government imposes a tax per pack of cigarettes, by how much will cigarette consumption fall? Show this in a diagram.
9. Given the following income elasticities of demand, would you classify the following goods as normal or inferior goods?
   a. Potatoes: elasticity = 0.5
   b. Pinto beans: elasticity = −0.1
   c. Bottled water: elasticity = 1.1
   d. Video cameras: elasticity = 1.4

10. Calculate the cross-price elasticity for the following goods. Are they substitutes or complements?
    a. The price of movie theater tickets goes up by 10 percent, causing the quantity demanded for video rentals to go up by 4 percent.
    b. The price of computers falls by 20 percent, causing the quantity demanded of software to increase by 15 percent.
    c. The price of apples falls by 5 percent, causing the quantity demanded of pears to fall by 5 percent.
    d. The price of ice cream falls by 6 percent, causing the quantity demanded of frozen yogurt to fall by 1 percent.

11. The following data on world coffee production from 1985 to 1987 come from the *CRB Commodity Year Book*.

	Production (millions of bags)	Price (dollars/pound)
1985	96	1.42
1986	80	2.01
1987	103	1.09

    a. Plot the observations on a scatter diagram with price on the vertical axis and production on the horizontal axis.
    b. Assume that production equals the quantity demanded around the world and that the coffee demand curve did not change between 1985 and 1987. First, calculate the price elasticity of demand for 1985 to 1986 using the midpoint formula. Next, calculate the same elasticities from 1986 to 1987. Would you say the demand for coffee is elastic or inelastic?
    c. If the two calculations in part (b) differ by much, explain why.

12. In 1992, the federal government placed a tax of 10 percent on goods like luxury automobiles and yachts. The boat-manufacturing industry had huge declines in orders for boats and laid off many workers, whereas the reaction in the auto industry was much milder. (The tax on yachts was subsequently removed.) Explain this situation using two supply and demand diagrams. Compare the elasticity of demand for luxury autos with that for yachts based on the experience with the luxury tax.

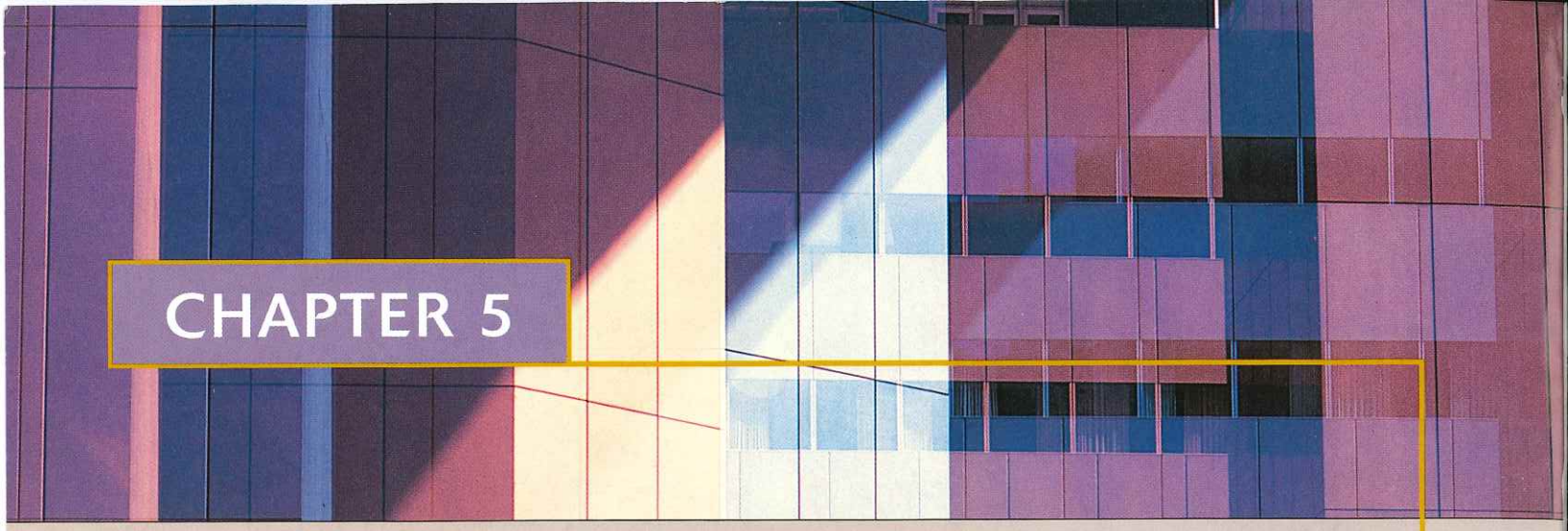
10. Calculate the cross-price elasticity for the following goods. Are they substitutes or complements?
    a. The price of movie theater tickets goes up by 10 percent, causing the quantity demanded for video rentals to go up by 4 percent.
    b. The price of computers falls by 20 percent, causing the quantity demanded of software to increase by 15 percent.
    c. The price of apples falls by 5 percent, causing the quantity demanded of pears to fall by 5 percent.
    d. The price of ice cream falls by 6 percent, causing the quantity demanded of frozen yogurt to fall by 1 percent.

11. The following data on world coffee production from 1985 to 1987 come from the *CRB Commodity Year Book*.

	Production (millions of bags)	Price (dollars/pound)
1985	96	1.42
1986	80	2.01
1987	103	1.09

    a. Plot the observations on a scatter diagram with price on the vertical axis and production on the horizontal axis.
    b. Assume that production equals the quantity demanded around the world and that the coffee demand curve did not change between 1985 and 1987. First, calculate the price elasticity of demand for 1985 to 1986 using the midpoint formula. Next, calculate the same elasticities from 1986 to 1987. Would you say the demand for coffee is elastic or inelastic?
    c. If the two calculations in part (b) differ by much, explain why.

12. In 1992, the federal government placed a tax of 10 percent on goods like luxury automobiles and yachts. The boat-manufacturing industry had huge declines in orders for boats and laid off many workers, whereas the reaction in the auto industry was much milder. (The tax on yachts was subsequently removed.) Explain this situation using two supply and demand diagrams. Compare the elasticity of demand for luxury autos with that for yachts based on the experience with the luxury tax.



CHAPTER 5

The Demand Curve and the Behavior of Consumers

This is a true story about a college professor who loves to teach introductory economics. The professor is younger than most college professors, but is hard of hearing and wears hearing aids in both ears. The professor teaches one of those large lecture courses, and most students aren't even aware that the professor wears the hearing aids.

In the middle of one of the lectures—in order to illustrate an important point about demand curves—the professor simultaneously brings one hand to one ear and the other hand to the other ear and suddenly pulls out both hearing aids, saying, "I can't hear a thing. If it were not for these hearing aids I wouldn't be here. I couldn't be a teacher. Do you know how much benefit I get from these hearing aids? Certainly more than from my car and maybe even more than from my house. If I had to give you a dollar amount, I would say that the benefit to me is about $60,000. Without the hearing aids, I would probably earn less, and I know my life would not be as enjoyable. Of course, I had to buy these hearing aids, and they are not very cheap. They cost me $500. But, you know, they cost me a lot less than they benefit me. The difference between $60,000 and $500 is $59,500, a huge amount. That difference is a measure of what the hearing aid market delivers to me over and above what I had to pay for the hearing aids. Most

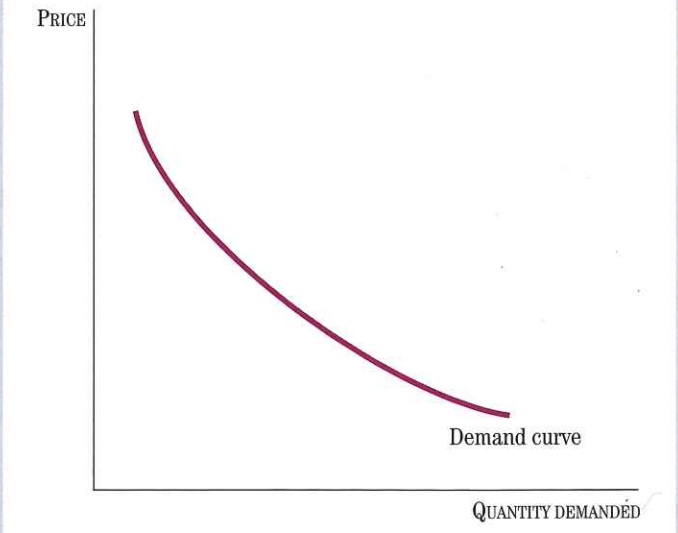


Figure 5.1
A Typical Demand Curve
Demand curves typically slope downward. In this chapter, we examine the behavior of the consumers who underlie the demand curve.

people would call that a good deal, but because I am an economics professor I call it a *consumer surplus.*"

In this chapter we show how and why the demand curve for any good—whether hearing aids, MP3 players, grapes, or bananas—can be used to measure the "good deal" or the "consumer surplus" that markets deliver to people.

Figure 5.1 shows a typical demand curve, with price on the vertical axis and quantity demanded on the horizontal axis. The demand curve is for an entire market, which might consist of millions of consumers. In this chapter we look under the surface of the demand curve and examine the behavior of these consumers. This examination has two purposes. The first is to see exactly what determines the slope and position of the demand curve—why it slopes downward and why changes in people's preferences or income cause it to shift.

The second purpose of examining the behavior of consumers is to see how well a market economy actually works. When we study the interactions of people in markets, it is the foundation of demand curves—people's preferences and choices—that we are investigating, not the demand curves themselves. Consumers do not go to the market with a demand curve; they go with certain preferences and objectives. One of the most important conclusions of the study of economics is that, under certain circumstances, a market economy works better than alternative systems to produce and allocate goods and resources. In order to understand that conclusion—to question it, to criticize it, to prove it, to defend it—we must look at the consumer behavior beneath the demand curve.

Our examination of consumer behavior in this chapter involves constructing a model. The main assumption of the model is that people make purposeful choices with limited resources to increase their satisfaction and better their lives. To make this assumption operational, economists have developed the idea of *utility*, which represents people's preferences for different items (products, jobs, leisure time) among a set of alternatives. We first define utility and then show how economists use utility to derive the slope and position of the demand curve.

Utility and Consumer Preferences

Every person has tastes and preferences for some goods relative to other goods. The millions of people who underlie a typical demand curve do not all have the same tastes and preferences, of course. Some people like Brussels sprouts; some people hate Brussels sprouts. We first focus on individual consumers and then show how the behavior of the millions of individuals adds up.

utility: a numerical indicator of a person's preferences in which higher levels of utility indicate a greater preference.

Utility is a numerical indicator of a person's preference for some goods compared to others. If one prefers some activity, such as eating a pizza and drinking two Cokes, to some other activity, such as eating two pizzas and drinking one Coke, then the utility from "one pizza and two Cokes" is greater than the utility from "two pizzas and one Coke." In general, if activity A is preferred to alternative B, then the utility from A is greater than the utility from B.

Be careful not to confuse the economist's definition of utility with the everyday meaning. If you look up *utility* in the dictionary, you will probably see the word *usefulness*, but to an economist, higher *utility* does not mean greater "usefulness"; it simply means that the item is preferred to another item. Watching "The Academy Awards" or "The NCAA Final Four" might give you more utility than attending a review session for your economics course, even though it is certainly not as useful for studying for the final.

A Consumer's Utility Depends on the Consumption of Goods

Let us consider an example of utility. Grapes are a product with which we have a lot of experience. They have been around for more than 4,000 years (at least since 2400 B.C. in ancient Egypt), and, in one form or another, they are still consumed around the world. Bananas are another popular fruit, also consumed around the world. Let us use grapes and bananas for our example of utility.

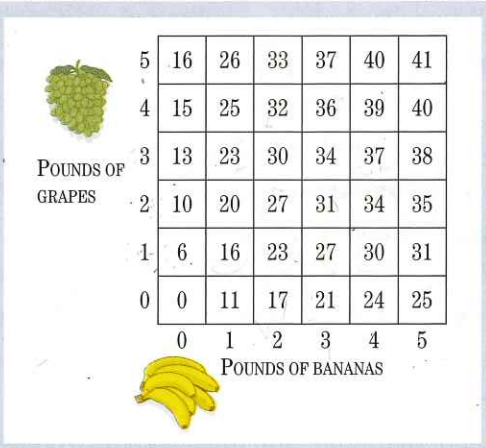


Figure 5.2
Example of Utility from Grapes and Bananas
The numbers inside the box give the utility from consuming the amounts of grapes and the amounts of bananas shown outside the box. For example, utility from 4 pounds of grapes and 3 pounds of bananas is 36. Combinations of grapes and bananas with higher utility are preferred to combinations with lower utility.

Figure 5.2 shows an example of the utility that one individual might get from consuming different amounts of grapes and bananas. Because every person is different, Figure 5.2 is just an example. You might imagine that the person is you, standing in front of a bin of fresh grapes and bananas in a grocery store, deciding how many pounds to buy. Or it could be the person in the picture at an open-air market who has some cash and is deciding how many pounds of grapes and bananas to buy with the cash.

Observe how Figure 5.2 is organized. The number of pounds of grapes is listed vertically on the left outside the box, from 0 up to 5 pounds. The number of pounds of bananas is listed horizontally below the box, from 0 over to 5 pounds. The entries inside the box show the utility from consuming different combinations of grapes and bananas.

To find the utility, first go to the row corresponding to the number of pounds of grapes and then go to the column corresponding to the number of pounds of bananas. The box at the intersection of the row and column shows the utility for the consumption of that specific combination of grapes and bananas. For example, if the individual consumes 2 pounds of grapes and 1 pound of bananas, then the utility is 20.

Note that this individual gets more utility from 2 pounds of grapes and 1 pound of bananas (utility = 20) than from 1 pound of grapes and 1 pound of bananas (utility = 16), which seems reasonable. In other

"If we buy two at full price, we can buy another six we don't want at half price."

words, 1 additional pound of grapes increases utility by 4 (from 16 to 20). This increase in utility is called marginal utility. In general, *marginal utility* is the increase in utility from consuming an additional unit of a good.

Utility Indicates Preference

We can now show exactly how utility is a numerical indicator of a person's preference for one good compared with another. Recall that when utility from one activity is greater than utility from an alternative, the activity is preferred to the alternative. According to Figure 5.2, the consumer prefers a combination of 4 pounds of grapes and 1 pound of bananas to a combination of 1 pound of grapes and 2 pounds of bananas because the utility of the former (25) is greater than the utility of the latter (23). Other combinations can be ranked similarly. In some cases there are ties; for example, the consumer is *indifferent* as to 2 pounds of grapes and 3 pounds of bananas versus 1 pound of grapes and 5 pounds of bananas because the utility of both is 31.

By ranking different combinations of goods in this way, we can see that a consumer's utility describes the consumer's preference for one good compared with another. Of all possible combinations, the one with the highest (maximum) utility is the one that is preferred to all the others. Thus, by maximizing utility, the consumer can be said to be making decisions that lead to the most preferred outcome from the viewpoint of the consumer. In this way, utility maximization implements the assumption that people make purposeful choices to increase their satisfaction.

An important fact about utility as exemplified in Figure 5.2 is that the units used to measure it do not matter. For example, suppose we multiply the utilities from grapes and from bananas in Figure 5.2 by 2, and then reexamine what utility implies about preferences. Rather than 23 units of utility, we would have 46 units of utility from 1 pound of grapes and 2 pounds of bananas. But rather than this combination, the consumer would still prefer 4 pounds of grapes and 1 pound of bananas, which would have a utility of $2 \times 25 = 50$. In fact, you can multiply utility by any positive number—even a billion or a billionth—and still get the same ordering of one combination compared to another.

The fact that the description of people's preferences does not depend on the units we use to measure utility is very important, because in reality economists have no way to measure utility. That is why Figure 5.2 does not give units. In particular, no one can say that one person's utility is higher or lower than another person's utility. The utilities of different people cannot be compared. An important feature of economists' use of utility is that it does not require or imply that the utilities of different people can be compared. Only the preference of a particular person for one type of good in comparison with another type of good is represented by utility.

REVIEW

- Utility is a numerical indicator of a person's preferences for different goods. For each combination of goods, there is a numerical value of utility.
- Combinations of goods with a higher utility are preferred to combinations of goods with a lower utility.
- The units by which utility is measured do not affect the preference for one combination compared with another.

The Budget Constraint and Utility Maximization

We have now shown how a consumer's preferences can be described by utility. Maximizing utility is an assumption equivalent to making purposeful choices to improve one's satisfaction. Now let us introduce the limits on the consumer's choice and explain how utility maximization works.

The Budget Constraint

budget constraint: an income limitation on a person's expenditure on goods and services.

Consumers are limited in how much they can spend when they choose between grapes or bananas or other goods. For example, suppose that the individual choosing between grapes and bananas is limited to spending a total of $8. That is, total spending on grapes plus bananas must be less than or equal to $8. This limit on total spending is called the **budget constraint.** In general, a budget constraint tells us that total expenditures on all goods and services must be less than a certain amount, perhaps the person's income for the year. The budget constraint is what limits the consumer's choices.

How much a consumer can spend and still remain within the budget constraint depends on the prices of the goods. For the example of grapes and bananas, if the consumer buys 1 pound of grapes at $2 and 2 pounds of bananas at $1, then expenditures are $4, well within the budget constraint of $8. But if 5 pounds of each were purchased at these prices, expenditures would be $15, a sum outside the budget constraint and, therefore, not possible.

Modeling a Consumer's Choice
This consumer, with a limited amount to spend, makes a choice that maximizes her utility. The combination of grapes and bananas that she prefers to other possible combinations of grapes and bananas must have a higher utility for her than the other combinations have.

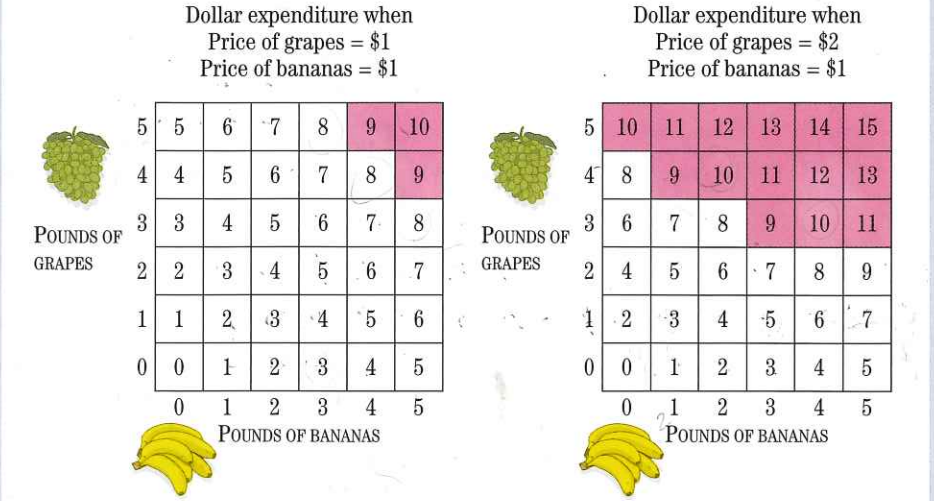


Figure 5.3
The Budget Constraint and Expenditures at Two Different Prices
The numbers inside the box give the total dollar expenditures on different combinations of grapes and bananas. For example, in the left box, where the price of both bananas and grapes is $1 per pound, the total dollar expenditure would be $7 for 4 pounds of grapes and 3 pounds of bananas. If the price of grapes were $2 per pound and the price of bananas were $1 per pound, as in the right box, that same combination would cost $11. The numbers in the red-shaded area are greater than the $8 budget constraint.

Figure 5.3 shows expenditures on grapes and bananas for two different situations. In both situations the price of bananas is $1; but in the situation in the left-hand box the price of grapes is $1, while in the right-hand box the price of grapes is $2. All the combinations of grapes and bananas from Figure 5.2 are shown in Figure 5.3. Several of the combinations are not within the $8 budget constraint; these are in the red-shaded area. Observe that when the price of grapes rises from $1 to $2, more combinations are outside the budget constraint and fewer are within the budget constraint. In general, a higher price for a good reduces consumption opportunities for the individual.

Maximizing Utility Subject to the Budget Constraint

utility maximization: an assumption that people try to achieve the highest level of utility given their budget constraint.

Given the utility in Figure 5.2 and the budget constraint in Figure 5.3, we can now show what happens when the individual maximizes utility subject to the budget constraint. **Utility maximization** means that people choose the highest possible level of utility given their budget constraint. In Figure 5.4, we show utility, from Figure 5.2, but now we shade in red the combinations for which expenditures are greater than the $8 budget constraint from Figure 5.3. The budget constraint means that you can't go in the red area.

Suppose first that the price of grapes is $1 per pound and the price of bananas is $1 per pound. By looking at the box on the left, you can find the highest level of utility achievable by the consumer with an $8 budget constraint. It is 39 units of utility from 4 pounds of grapes and 4 pounds of bananas. These values are shown in blue boldface type. This is the most preferred combination that the individual can buy

Maximum utility subject to budget constraint when Price of grapes = $1 Price of bananas = $1

Pounds of grapes \ Pounds of bananas	0	1	2	3	4	5
5	16	26	33	37	40	41
4	15	25	32	36	**39**	40
3	13	23	30	34	37	38
2	10	20	27	31	34	35
1	6	16	23	27	30	31
0	0	11	17	21	24	25

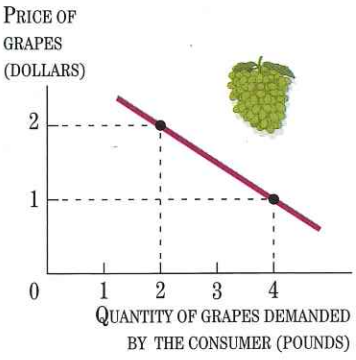
Maximum utility subject to budget constraint when Price of grapes = $2 Price of bananas = $1

Pounds of grapes \ Pounds of bananas	0	1	2	3	4	5
5	16	26	33	37	40	41
4	15	25	32	36	39	40
3	13	23	30	34	37	38
2	10	20	27	31	**34**	35
1	6	16	23	27	30	31
0	0	11	17	21	24	25

Figure 5.4
Maximizing Utility Subject to the Budget Constraint at Two Different Prices
The red-shaded areas are the same as the red-shaded areas in Figure 5.3 and therefore represent combinations for which total expenditures would be greater than the $8 budget constraint. They are not possible choices. The bold number is the maximum level of utility for which spending is less than or equal to $8. In the left box, the maximum utility of 39 represents a choice of 4 pounds of grapes and 4 pounds of bananas. In the right box, with the higher price of grapes, the maximum utility is 34, corresponding to 2 pounds of grapes and 4 pounds of bananas. Hence, a higher price of grapes leads to a lower quantity of grapes demanded.

and still remain within the budget constraint. Thus, utility maximization predicts that the consumer will purchase 4 pounds of grapes and 4 pounds of bananas at these prices. There is no other combination that will yield greater utility and still be within the $8 budget constraint.

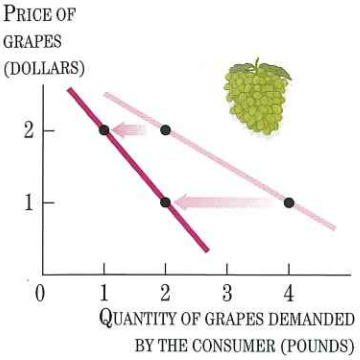
Price of Grapes	Quantity of Grapes Demanded by the Consumer
$1	4 pounds
$2	2 pounds



■ **Effect of a Change in Price: A Movement Along a Demand Curve.** Now suppose that the price of grapes rises from $1 to $2 per pound with the price of bananas staying the same. The options for expenditures in this case are shown in the box on the right in Figure 5.4. You can see the combination—2 pounds of grapes and 4 pounds of bananas—for which the maximum utility (34) is reached.

Now observe something very important: The quantity of grapes demanded at the higher price of grapes is less than the quantity demanded at the lower price of grapes. When the price of grapes is $1 per pound, 4 pounds of grapes are purchased; at the price of $2 per pound, 2 pounds of grapes are purchased. Thus, we have shown that the assumption that people maximize utility subject to a budget constraint implies that a higher price leads to a reduced quantity demanded. In other words, as shown in the margin, we have derived two points on a demand curve for the consumer. When the price goes up, the quantity demanded goes down; when the price goes down, the quantity demanded goes up. These are *movements along* a demand curve, as shown in the small graph in the margin, which indicates the two prices and the quantity demanded at each price; a line is drawn through the points to illustrate the downward-sloping demand curve.

Price of Grapes	Quantity of Grapes Demanded by the Consumer
$1	2 pounds
$2	1 pound



■ **Effect of a Change in Income: A Shift in the Demand Curve.** Now consider the effect of a change in the individual's income on the quantity of grapes that the individual will purchase. Suppose that the individual now has $5 to spend rather than $8; in other words, there is a $3 reduction in the individual's income. What will happen to the quantity of grapes demanded by the consumer? If the price is again $1 per pound of grapes and $1 per pound of bananas, then the boxes on the left of Figures 5.3 and 5.4 apply. (You might want to shade in the additional area of these two boxes that represents combinations of grapes and bananas that are not feasible with only $5 to spend.) If expenditures are limited to $5, then the maximum utility (31) occurs when 2 pounds of grapes and 3 pounds of bananas are purchased. Recall that with $8 to spend, the consumer was willing to buy 4 pounds of grapes. Thus, at a price of grapes of $1 per pound, the decrease in income leads to a reduction in the amount of grapes the consumer is willing to buy.

We can also calculate the effects of the income change at a different set of prices. If the price is $2 for grapes and $1 for bananas, then the boxes on the right of Figures 5.3 and 5.4 apply. (Again, you might want to shade in the additional area of these two boxes that represents combinations of grapes and bananas that are not feasible with only $5 to spend.) With a limit of $5 to spend, the maximum utility (27) occurs when 1 pound of grapes and 3 pounds of bananas are purchased. The decrease in income leads to a decrease in the amount of grapes the consumer will purchase at each price of grapes. Thus, as shown in the margin, we have derived another demand curve corresponding to the decreased amount of income. Observe that the demand curve with the lower amount of income ($5) is shifted to the left compared with the demand curve with the higher amount of income ($8). Thus, we have shown explicitly that a change in income will *shift* the consumer's demand curve.

■ **Income and Substitution Effects of a Price Change.** Using the concepts of utility and the budget constraint, economists distinguish between two separate reasons why an increase in the price leads to a decrease in the quantity demanded. These are (1) the income effect and (2) the substitution effect.

The Income Effect of a Change in the Price We noted how an increase in the price reduces the number of options available to the consumer. When the price of grapes rises from $1 to $2 per pound, choices such as 4 pounds of grapes and 4 pounds of bananas are no longer within the budget constraint, although they were within the budget constraint at a grape price of $1 per pound. A total of fourteen options are outside the budget constraint (in the red-shaded area of the right tables of Figures 5.3 and 5.4) when the price of grapes is $2, whereas only three are outside the budget constraint (in the red-shaded area of the left tables of Figures 5.3 and 5.4) when the price is $1. This reduction in the options when the price rises is similar to what would happen if the consumer suddenly had less income to spend on both goods without any change in prices. Let's show this effect in the context of the banana and grape example.

Originally we observed a price of grapes and bananas of $1 per pound, and $8 of available income, with an optimal quantity of 4 pounds of grapes and 4 pounds of bananas, and 39 units of utility. After the increase in the price of grapes to $2 per pound, the consumer now purchases 2 pounds of grapes and 4 pounds of bananas, with a lower level of utility of 34 units. Now imagine that we want to achieve the same lower level of utility (34 units) in a different way, by keeping the original prices constant while lowering the available income. If prices are still $1 per pound for grapes and bananas, Figure 5.5 shows that a budget constraint of $6 eliminates 10 combinations of grapes and bananas, and the consumer buys 3 pounds of grapes and 3 pounds of bananas to maximize the utility level at 34 units. (Note that in this

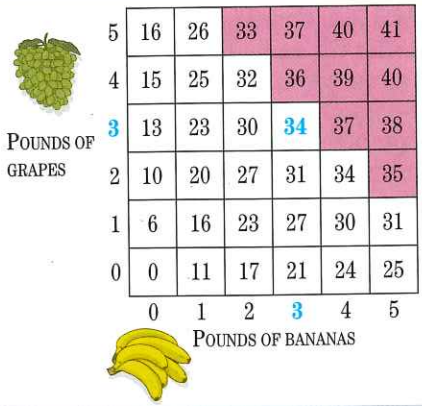


Figure 5.5
Income Effect
The red-shaded area represents combinations that are not feasible for a price per pound of $1 for both grapes and bananas, and an available income of $6. The consumer may choose to buy 3 pounds of grapes and 3 pounds of bananas, showing the effect of a reduction of income on the quantity consumed.

particular example, 2 pounds of grapes and 4 pounds of bananas also yield 34 units of utility and the consumer is indifferent between two combinations of goods. For a more general analysis, read the appendix to this chapter.) The decrease from 4 pounds of grapes to 3 pounds of grapes measures the income effect.

The **income effect** is the amount by which quantity demanded falls because of the decline in real income from the price increase. Of course, a reduction in the grape price will have the opposite effect: It will increase the real income the consumer has to spend on both goods and in particular on grapes. The income effect is a general phenomenon that applies to all normal goods; for example, when the price of gasoline rises, people will spend less on gasoline in part because their real income has declined. With less real income, they will spend less on most goods and services.

The Substitution Effect of a Change in the Price An increase in the price of grapes with no change in the price of other goods causes an increase in the relative price of grapes. Because grapes become relatively more expensive, people will switch their purchases away from grapes toward other goods even if there is no income effect. The **substitution effect** is the amount by which the quantity demanded falls when the price rises, exclusive of the income effect.

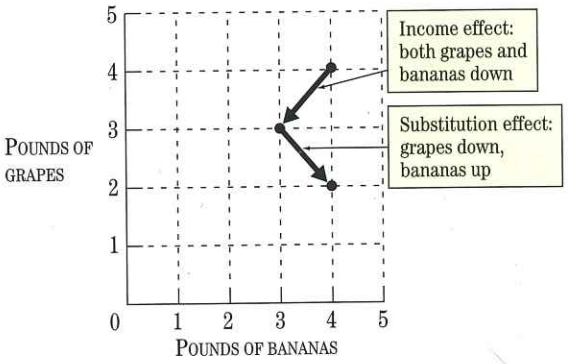
income effect: the amount by which the quantity demanded falls because of the decline in real income from a price increase.

substitution effect: the amount by which quantity demanded falls when the price rises, exclusive of the income effect.

In the grape-banana example, the increase in price from $1 to $2, with $8 of available income, makes the combination of 3 pounds of grapes and 3 pounds of bananas of Figure 5.5 unavailable too. The feasible combination that maximizes utility is 2 pounds of grapes and 4 pounds of bananas, with a utility of 34 units. The total reduction in grape consumption of 2 pounds can be split, in this example, into 1 pound (from 4 to 3 pounds) due to the income effect and 1 pound (from 3 to 2 pounds) due to the substitution effect. The income effect reduces both grape and banana consumption, while the substitution effect just reduces the quantity demanded of the good with the increasing price, as illustrated in the table and graph below. The exact size of the income and substitution effects differs from example to example. In some cases the income effect is larger than the substitution effect; in other cases the substitution effect is larger.

	Quantity of Grapes		Quantity of Bananas
Original quantity demanded	4		4
	↓	income effect	↓
	3		3
	↓	substitution effect	↓
New quantity demanded	2		4

Visualizing the Income and Substitution Effects
Plotting the quantities from the above table in a graph with one good on the vertical axis and another on the horizontal axis helps in visualizing the income and substitution effects.



REVIEW

- The demand curve can be derived from the assumption that people maximize utility subject to a budget constraint.
- A change in the price of a good changes the number of combinations within the budget constraint and thus changes the utility-maximizing quantity demanded. This represents a movement along the demand curve.
- A change in income also changes the budget constraint and thus also changes the utility-maximizing quantity demanded. This represents a shift of the demand curve.
- The total effect of a change in the price on the quantity demanded can be divided into two parts—an income effect and a substitution effect.

Willingness to Pay and the Demand Curve

The choice between one good (grapes) and another good (bananas) in the previous section is useful for showing how to derive a demand curve from the basic idea that consumers maximize utility subject to a budget constraint. Now we want to use the demand curve to determine how well markets work for consumers. This requires moving beyond the simple choice between one good and another good and considering the choice between one good and all possible other goods.

Measuring Willingness to Pay and Marginal Benefit

Table 5.1
Willingness to Pay (Benefit) and Marginal Benefit

Quantity of *X*	Willingness to Pay for *X* (Benefit from *X*)	Marginal Benefit from *X*
0	$0.00	—
1	$5.00	$5.00
2	$8.00	$3.00
3	$9.50	$1.50
4	$10.50	$1.00
5	$11.00	$.50

The connecting lines emphasize how marginal benefit is the *change* in benefit (or willingness to pay) as one more unit of a good is consumed.

marginal benefit: the increase in the benefit from, or the willingness to pay for, one more unit of a good.

Suppose we asked a person who is consuming a zero amount of good *X*, "How much money would you be willing to pay for one unit of *X*?" Because the money that the person would pay can be used to buy all goods, not just one good, the question implicitly asks the person to compare *X* with all other goods. In general, the answer to this question would depend on how much utility would increase with one unit of *X* and on how much utility would decrease because less would be spent on other goods, given the budget constraint. In other words, the answer would depend on the person's preferences for *X* and all other goods as represented by utility.

Suppose that the answer is $5. Let us assume that the answer to the question gives us the true measure of the consumer's preferences. Then, once we get an answer to the first question, we could ask, "How much would you be willing to pay for two units of *X*?" Suppose the answer is, "I would be willing to pay $8." We could then continue to ask the consumer about more and more units of *X*. We summarize the hypothetical answers in Table 5.1. The column labeled "Willingness to Pay for *X*" tabulates the answers to the question.

Assuming that the answers to the questions are true, willingness to pay measures how much the consumer would *benefit* from different amounts of *X*. **Marginal benefit** from *X* is the increase in benefit from, or the willingness to pay for, one more unit of a good. As a person consumes more and more of a good, the marginal benefit from additional amounts is likely to diminish. Imagine you are very hungry, there is no food in the house, and you are craving pizza. At this point you might be willing to pay $5.00 for a big, hot slice of pizza. Now suppose you already have that slice of pizza, but you're still a little hungry. You might be willing to pay $3.00 for an additional slice. The more pizza you have, the less you are willing to pay for even more pizza. Observe that the marginal benefit in Table 5.1 diminishes as more is consumed.

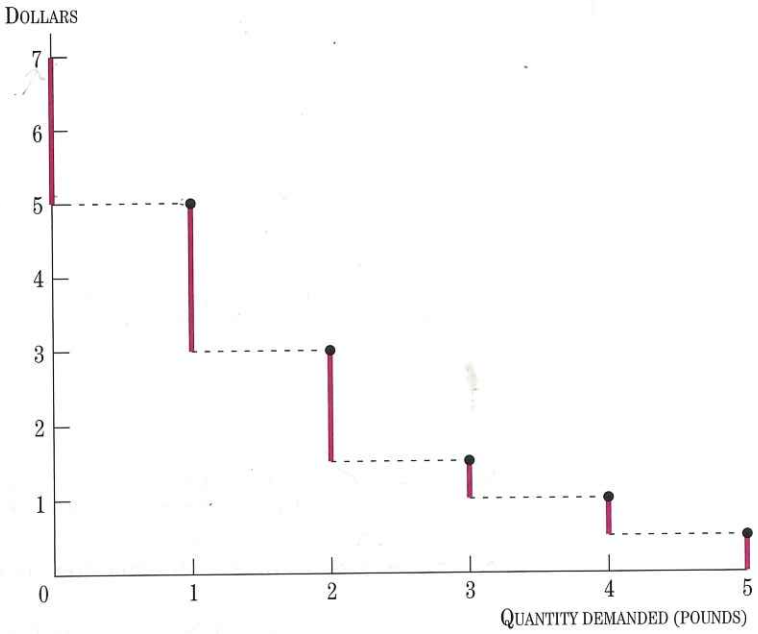


Figure 5.6
Derivation of the Individual Demand Curve
The dots are exactly the same as the marginal benefit in Table 5.1. At each dot, price equals marginal benefit. The vertical lines indicate how much is demanded at each price if the consumer is restricted to purchasing whole pounds.

Graphical Derivation of the Individual Demand Curve

A demand curve can be derived from the information about willingness to pay (benefit) and marginal benefit of *X* for the person described in Table 5.1. Suppose that *X* is raisins (rice, salt, tea, orange juice, CDs, movies, or any other good will serve just as well as an example). Suppose that the person has $10 to spend on raisins and other goods. We want to ask how many pounds of raisins the person would buy at different prices. We imagine different hypothetical prices for raisins, from astronomical levels like $7 a pound to bargain basement levels like $.50 a pound.

To proceed graphically, we first plot the marginal benefit from Table 5.1 in Figure 5.6. Focus first on the black dots in Figure 5.6. The lines will be explained in the next few paragraphs. The horizontal axis in Figure 5.6 measures the quantity of raisins. On the vertical axis we want to indicate the price as well as the marginal benefit, so we measure the scale of the vertical axis in dollars. The black dots in Figure 5.6 represent the marginal benefit an individual gets from consuming different amounts of raisins.

How many pounds of raisins would this person consume at different prices for raisins? First, suppose that the price is very high—$7 a pound. Draw an arrow pointing to this $7 price in Figure 5.6. We are going to derive a demand curve for this individual by gradually lowering the price from this high value and seeing how many pounds would be purchased at each price. As the price declines, you can slide your arrow down the vertical axis. For each price, we ask the same question: How many pounds would the person buy? To make things simple at the start, assume that the person buys only whole pounds of raisins. You might want to imagine that the raisins come in 1-pound cellophane packages. We consider fractions of pounds later.

Suppose, then, that the price is \$7 a pound. The marginal benefit from 1 pound of raisins is \$5. Thus, the price is greater than the marginal benefit. Would the person buy a pound of raisins at this price? Because the price the person would have to pay is greater than the marginal benefit, *the answer would be no;* the person would not buy a pound of raisins at a price of \$7. If the minimum amount of raisins that can be purchased is 1 pound, then the person will buy no pounds at a price of \$7 per pound. We have shown, therefore, that the quantity demanded of raisins is *zero* when the price is \$7.

Check the answer. To check that zero is the correct answer, suppose that the person did buy 1 pound of raisins for \$7. Then the person would have \$3 left over (\$10 − \$7 = \$3), and with the benefit of 1 pound of raisins being \$5, that would be a total of \$8 (since \$3 + \$5 = \$8). Because \$8 is less than the original \$10, by buying 1 pound of raisins the person would be worse off than by buying no raisins.

Now start to lower the price. As long as the price is more than \$5, the person will not buy any raisins. Hence, the quantity demanded at all prices higher than \$5 is zero. We indicate this by the red line on the vertical axis above the \$5 mark.

Now watch what happens when the price drops to \$5. The marginal benefit from a pound of raisins is \$5 and the price is \$5, so the marginal benefit of the pound of raisins is sufficient to cover the price. Now the person gets sufficient marginal benefit to buy 1 pound of raisins. Strictly speaking, the price would have to be a little less than \$5, perhaps \$4.99, for the person to get more by purchasing 1 rather than 0 pounds of raisins. At a price of \$5, the person will be indifferent between 1 pound and 0 pounds. At a price of \$4.99, the person would buy, but at a price of \$5.01, the person would not buy. The price of \$5 is right in between. However, let us assume that the person buys 1 pound rather than 0 pounds when the price is \$5. This is indicated in Figure 5.6 by showing that the quantity demanded is given by the black dot at 1 pound when the price equals \$5.

Check the answer. To check that 1 pound is the correct answer, look at how well off the person is after a decision to buy 1 pound compared with a decision to buy 2 pounds. If the person buys 1 pound at \$4, then the person has \$6 left over (\$10 − \$4 = \$6), and with the benefit (willingness to pay) being \$5 for that pound, the person would be well off in the amount of \$11 (since \$6 + \$5 = \$11). Now, if the person buys 2 pounds at \$4 per pound, then the person has \$2 left over (\$10 − \$8 = \$2), and with the benefit (willingness to pay) being \$8 for the 2 pounds, the person has the equivalent of \$10 (since \$2 + \$8 = \$10). Because \$11 is greater than \$10, the person is better off buying 1 pound than buying 2 pounds.

Continue lowering the price, slipping the arrow down the axis. The quantity demanded will stay at 1 pound as long as the price remains above the marginal benefit of buying another pound of raisins, or \$3. We therefore extend the red line downward at 1 pound as the price falls from \$5 down to \$3. Consider, for example, a price of \$4. The person has already decided that 1 pound will be bought; the question is whether a second pound of raisins is worthwhile. Another pound has a marginal benefit of \$3 (willingness to pay goes from \$5 to \$8 as the quantity increases from 1 to 2 pounds). The person has to pay \$4, which is more than the marginal benefit. Hence, *the quantity demanded stays at 1 pound when the price is \$4.* However, when the price falls to \$3, another pound is purchased. That is, when the price is \$3, the quantity demanded is 2 pounds, which is shown graphically by the black dot at 2 pounds.

Now suppose the price falls below \$3, perhaps to \$2. Is a third pound purchased? The marginal benefit of a third pound is \$1.50; is it worth it to buy a third pound at \$2 per pound? No. The quantity demanded stays at 2 pounds when the price is between \$3 and \$1.50, which we denote by extending the red line downward from the black dot at 2 pounds. This story can be continued. As the price continues to fall, more pounds of raisins are demanded.

By considering various prices from over \$5 to under \$.50, we have traced out an **individual demand curve** that slopes downward. As the price is lowered, more raisins are purchased. The demand curve is downward-sloping because of diminishing marginal benefit. At each black dot in the diagram, price equals the marginal benefit.

individual demand curve: a curve showing the relationship between quantity demanded of a good by an individual and the price of the good.

The jagged shape of the demand curve in Figure 5.6 may look strange. It is due to the assumption that only 1-pound packages of raisins are considered by the consumer. In the case of raisins, it is usually possible to buy fractions of a pound, and if the marginal benefits of the fractions are between the values for the whole pounds, then the demand curve will be a smooth line, as shown in Figure 5.7. Then price will equal marginal benefit not only at the black dots but also on the lines connecting the dots. If you are unsure of this, imagine creating a new Table 5.1 and Figure 5.6 with *ounces* of raisins. There will be a point at each ounce, and with 16 ounces per pound, there will be so many points that the curve will be as smooth as Figure 5.7.

READING THE NEWS ABOUT... What Neuroscience Can Tell Us About Economic Behavior

Our study of utility maximization in this chapter assumes that people try to maximize their utility, a numerical value representing a person's tastes and preferences, given a certain budget constraint. While we are often faced with similar cut-and-dried choices in real life (should I get the soda or the iced tea?), our choices more often reflect a variety of preferences that are hard to quantify and place a specific value on. As a result, our decision-making in real life may not always seem completely "rational." We may make economic decisions that will give us a short-term gain but a long-term loss. We may give our money away—because charitable giving may make us feel beneficent or because it seems like the right thing to do.

In the article that follows, you'll read about the work going on in the area of behavioral economics, or neuroeconomics, which looks more closely at the ways people make decisions and how they go about "satisfying their preferences."

Mind games

From the *Economist*, print edition
January 13, 2005

Can studying the human brain revolutionise economics?

ALTHOUGH Plato compared the human soul to a chariot pulled by the two horses of reason and emotion, modern economics has mostly been a one-horse show. It has been obsessed with reason. In decisions from how much to produce to whether to save and invest, humans have been assumed to be coolly rational calculators of their own self-interest. Over the past few years, however, evidence from psychology has persuaded many economists that reason does not always have its way. Now, judging from a series of presentations at the American Economic Association meetings in Philadelphia last weekend, a burgeoning new field dubbed "neuroeconomics" seems poised to provide fresh insights on how the two horses together produce economic behaviour.

The current bout of research is made possible by the arrival of new technologies such as functional magnetic-resonance imaging, which allows second-by-second observation of brain activity. At several American universities, economists and their collaborators in the neurosciences have been placing human subjects in such brain scanners and asking them to perform a variety of economic tasks and games.

For example, the idea that humans compute the "expected value" of future events is central to many economic models. Whether people will invest in shares or buy insurance depends on how they estimate the odds of future events weighted by the gains and losses in each case. Your pension, for example, might have a very low expected value if there is a large probability that bonds and shares will plunge just before you retire.

Brian Knutson, of Stanford University, carried out one recent brain-scan experiment to understand how humans compute such things. Subjects were asked to perform a task, in this case pressing a button during a short interval in which a certain shape was flashed on to a screen. In some trials, the subjects could win up to $5 if successful, while in others they would have to defend against a $5 loss. Before presenting the target, the researchers signalled to subjects which kind of trial they were in.

Brain activity in certain neural systems seemed to reveal a strong correlation with the amount of money at stake. Moreover, the prospects of gains and losses activated different parts of the brain. Traditional economists had long thought—or assumed—that the prospect of a $1,000 gain could compensate you for an equally likely loss of the same size. In subsequent trials, subjects were given another signal: one that provided an estimate of the odds of success. That allowed the researchers to identify the regions of the brain used for recognising an amount of money and for estimating the probability of winning (or losing) it. Having identified these regions, the hope is that future work can measure how the brain performs in situations such as share selection, gambling or deciding to participate in a pension scheme.

David Laibson, an economist at Harvard University, thinks that such experiments underscore the big role that expectations play in a person's well-being. Economists have usually assumed that people's well-being, or "utility", depends on their level of consumption, but it might be that changes in consumption, especially unexpected downward ones, as in these experiments, can be especially unpleasant.

Mr. Laibson's own work tries to solve a different riddle: why people seem to apply vastly different discount rates to immediate and short-term rewards compared with rewards occurring well into the future. People tend much to prefer, say, $100 now to $115 next week, but they are indifferent between $100 a year from now and $115 in a year and a week. In one recent experiment, noted in our science section on October 30th, Mr. Laibson and others found that the brain's response to short-term riches (in this case, gift certificates of $15 or $20) occurs largely in the limbic system, a region that governs emotion. By contrast, the prospect of rewards farther into the future triggers the prefrontal cortex, which is often associated with reason and calculation. Thus, choosing immediate economic gratification, by spending excessively on credit cards or not saving enough even though you "know better", could be a sign that the limbic system is in charge. Government policies, such as forced savings or "cooling off" periods for buying property or cars, may be one remedy.

And then there is trust and deception. Colin Camerer, of the California Institute of Technology, has conducted experiments in which brain-scanned participants play strategic games with anonymous partners. When players are doing the best that they can to "win" the game by anticipating their opponents' moves, their brains tend to show a high degree of coordination between the "thinking" and the "feeling" regions. Economic equilibrium, by this measure, is an identifiable "state of mind".

Don't let it go to your head

Some neuroeconomists claim that such brain-scanning experiments are the start of a revolution in economics. No longer will economists rely on crude statistical models of how people behave in response to a policy change, such as an interest-rate rise or a tax increase. Instead, they will be able to peer directly into the brain to predict behaviour.

One day, perhaps; but much work remains. Identifying the parts of the brain that control economic actions is one thing. Harder tasks include determining how neural systems work together to create behaviour, and how wide is the variation in brain patterns between different people. Then there are age-old questions of free will: is your failure to save for old age simply a lifestyle choice, or is it down to faulty brain circuits? Neuroeconomics is already providing fascinating conclusions. But Plato's chariot will remain an alluring explanation for a while yet.

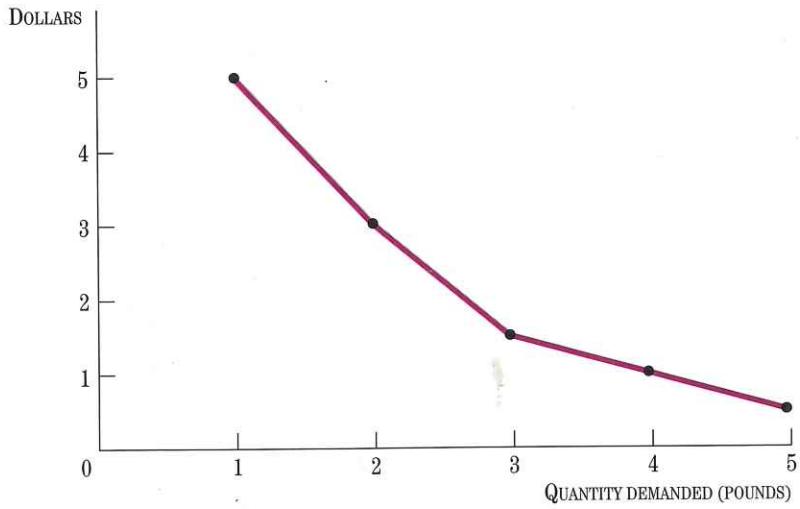


Figure 5.7
A Smooth Individual Demand Curve
If the consumer can buy fractions of a pound and if the marginal benefits of these fractions are between the whole-pound amounts, the demand curve becomes a smooth line, as in the figure, rather than the series of steps in Figure 5.5. In some cases, such as the demand for cars, we cannot consider fractions, and so these individual demand curves will look like steps.

The Price Equals Marginal Benefit Rule

We have discovered an important principle of consumer behavior. If the consumer can adjust consumption of a good in small increments—such as fractions of a pound—then the consumer will buy an amount for which the *price equals marginal benefit*. This condition can be applied to any good—apples, peanuts, comic books, the number of movies you see each year—not just raisins.

The price equals marginal benefit rule can explain a number of otherwise puzzling observations. For example, consider Adam Smith's diamond-water paradox. As Smith put it, "Nothing is more useful than water: but it will purchase scarce any thing; scarce any thing can be had in exchange for it. A diamond, on the contrary, has scarce any value in use; but a very great quantity of other goods may frequently be had in exchange for it."[1] Why are diamonds expensive and water cheap even though diamonds are less "useful" to the world's population than water?

The price equals marginal benefit rule helps explain the paradox. The price of diamonds will be high if the marginal benefit of diamonds is high. The price of water will be low if the marginal benefit of water is low. As we saw earlier, the marginal benefit of something declines the more people consume of it. Thus, water has relatively low marginal benefit because with water being so plentiful, people consume much of it every day. The marginal benefit is low even though the total benefit from water consumption in the world is very high. On the other hand, diamonds have a high marginal benefit because with diamonds being so scarce, people consume relatively little of them. The marginal benefit of diamonds is high even though the total benefit of diamonds may be low. Thus, the price equals marginal benefit rule explains the diamond-water paradox.

1. Adam Smith, *Wealth of Nations* (New York: Modern Library Edition, 1994), pp. 31–32.

REVIEW

- People's preferences are reflected in their willingness to pay for different amounts of a good. Because dollars can be used to buy any good, willingness to pay compares one good with all other goods.
- The marginal benefit from a good is the increase in the benefit from, or the willingness to pay for, one additional unit of a good.
- Consumers choose an amount of a good that makes the marginal benefit equal to the price.
- The individual demand curve slopes downward because of diminishing marginal benefits from consuming more of a good. As the price falls, the individual will choose a quantity with a lower marginal benefit. Hence, the quantity demanded goes up as the price falls.

The Market Demand Curve

market demand curve: the horizontal summation of all the individual demand curves for a good; also simply called the demand curve.

Thus far, we have graphically derived the *individual* demand curve. Now we consider the **market demand curve,** which is the sum of the individual demand curves. Figure 5.8 shows how we do the summing up. The figure shows the demand curves for raisins for two individuals, Ann and Pete. To get the market demand curve, add up, for each given price, the total amount demanded by both Pete and Ann. For example, at a price of $5, Pete's demand is 1 pound and Ann's demand is 1 pound. The market demand is then 2 pounds. When the price is $3 a pound, the demand is 2 pounds for Pete and 2 pounds for Ann, or 4 pounds for the market as a whole. Obviously, the market for raisins consists of more than just Pete and Ann. To get the whole market, you would have to sum up the demands for millions of people.

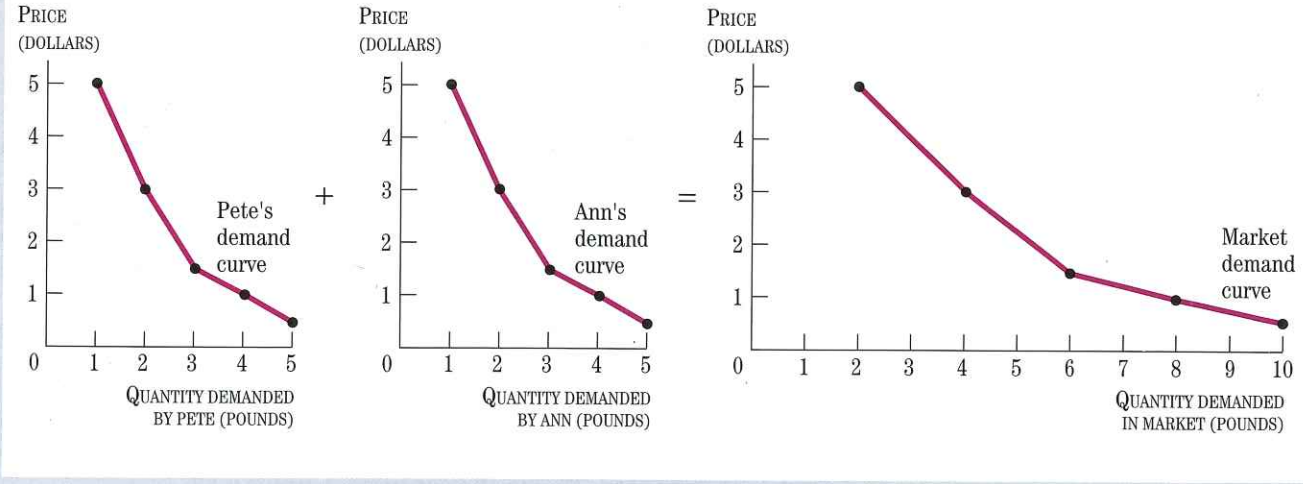


Figure 5.8 Derivation of the Market Demand Curve

The market demand curve is the sum of the demand curves of many individuals. The figure shows this for only two individuals with the same preferences. As more individuals with a diversity of tastes are added, the market demand curve becomes smoother and looks more like Figure 5.1.

Different Types of Individuals

In Figure 5.8, Pete's and Ann's demand curves are the same. They do not have to be. In fact, it is most likely that Ann and Pete have different preferences. Pete could be a peanut fan and be willing to pay less for raisins than Ann. It is incorrect to assume that everyone will be willing to pay the same amount for any good. There are all kinds of people in the world with different preferences. But you can still add up the demands of all these people at any given price to get the market demand curve. As you add up many individual demand curves for different types of people, the market demand curve gets smoother, even when the product cannot be bought in fractions of a unit. For example, the market demand curve for cars is smooth even though most individuals buy either zero, one, or perhaps two cars. When you add the demand curves for millions of people, the market demand curve for cars looks like the market demand curve (Figure 5.1) that we typically draw—smooth and downward-sloping. To confirm your understanding of the market demand curve, make sure you work through problems 5 and 7 at the end of the chapter.

REVIEW

- The market demand curve is derived from individual demand curves. At each price we add up how much is demanded by all individuals; the total is the market demand at that price.
- Even if the individual demand curves are not smooth, the market demand curve will be smooth because people have different tastes and preferences and prefer different benefits.

Consumer Surplus

consumer surplus: the difference between what a person is willing to pay for an additional unit of a good—the marginal benefit—and the market price of the good; for the market as a whole, it is the sum of all the individual consumer surpluses, or the area below the market demand curve and above the market price.

In many cases, people are willing to pay more for an item consumed than they have to pay for it. You might be willing to pay five times the $8 admission price to see your favorite movie. But like everyone else in line, you pay only $8, even though it is worth $40 to you. The difference between the $40 and the $8 is called *consumer surplus*.

In general, **consumer surplus** is the difference between the willingness to pay for an additional item (say, $40 for a movie)—its marginal benefit—and the price paid for it (say, $8 for a movie). Suppose the price of raisins is $4 per pound. Then the consumer in our previous example purchases 1 pound and the marginal benefit of that pound is $5. In that situation, the consumer gets a consumer surplus because the marginal benefit of the raisins to the consumer is $5 but the price paid is only $4. Consumer surplus is the difference, or $1. If the price were $3.50, then the consumer surplus would be greater, or $1.50.

Suppose the price falls further, so that two items are purchased. Consumer surplus is then defined as the sum of the differences between the marginal benefit of each item and the price paid for the item. For example, if the price per pound of raisins is $2, as in Figure 5.9, then 2 pounds of raisins will be purchased and the consumer surplus will be $5 − $2 = $3 for the first pound plus $3 − $2 = $1 for the second pound, for a total of $4. That is, the consumer surplus is $4.

Figure 5.9 shows graphically how consumer surplus is the area between the demand curve and the line indicating the price. In Figure 5.9, the total shaded area is equal to 4, consisting of two rectangular blocks, one with an area of 3 and the other

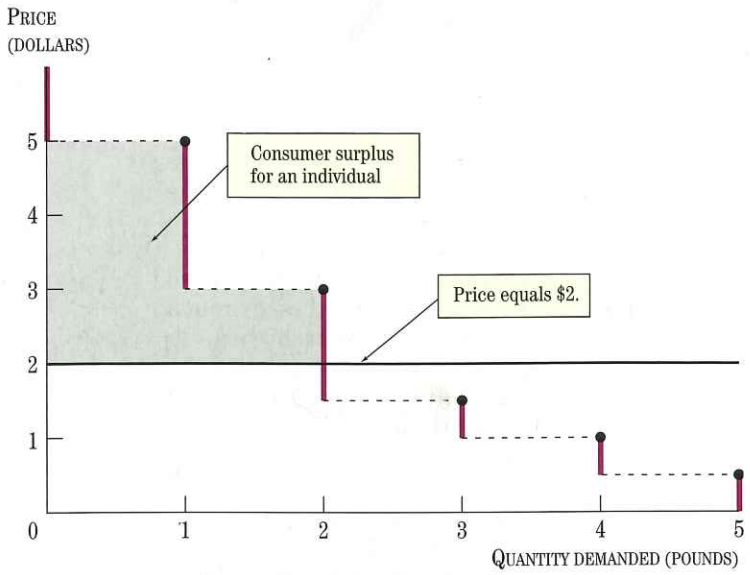


Figure 5.9
Consumer Surplus for an Individual
The consumer surplus is the difference between the marginal benefit a person gets from consuming a good and the price. It is given by the area between the demand curve and the price.

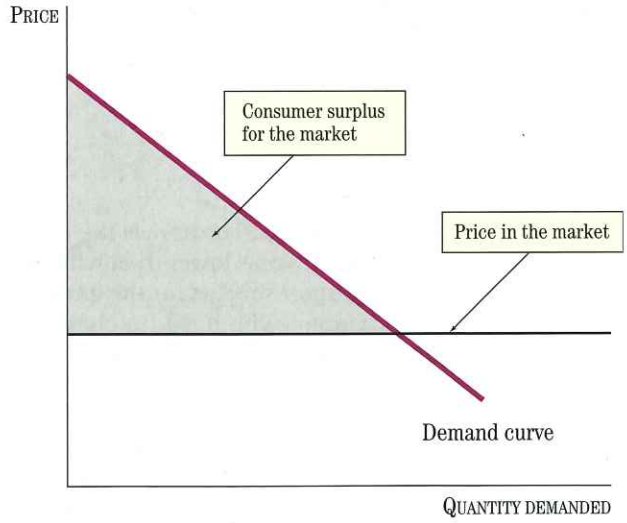


Figure 5.10
Consumer Surplus for the Market
The sum of the consumer surplus for all individuals in the market is the area between the demand curve and the price.

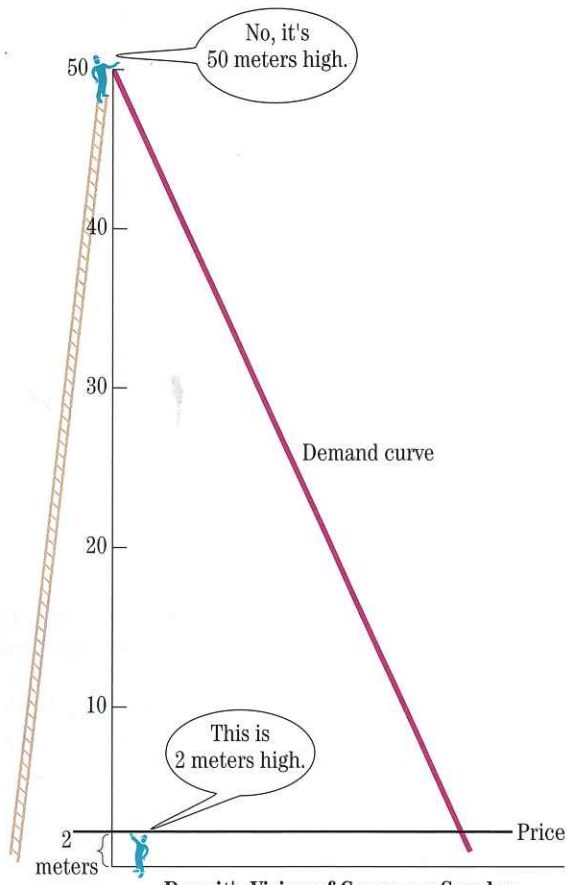
Building Roads and Bridges with Consumer Surplus

You may be surprised to learn that the concept of consumer surplus was not invented by an economist but by an engineer. Jules Dupuit was a civil engineer living in France in the mid-1800s. He wanted to demonstrate that the value of the roads and bridges he was building was much more than what people were willing to pay to use them. Consumer surplus was his idea of a demonstration. If a person paid a toll to cross a bridge, then the price of the toll could be as much of an underestimate of the benefit as the price of hearing aids is to the professor of economics in the introduction to this chapter. Thus, his consumer surplus argument helped persuade people to build more bridges and roads.

Dupuit offered a visual description of consumer surplus: "If society is paying 500 million for the services rendered by the road, that only proves one thing—that [the benefit from the road] is at least 500 million. But it may be a hundred times or a thousand times greater. . . . If you take the [500 million] as the figure . . . you are acting like a man who, wishing to measure the height of a wall in the dark and finding that he cannot reach the top with this raised arm says: 'This wall is two meters high, for if it were not, my hand would reach above it.' In daylight and equipped with a ladder . . . our alleged two-meter wall is fifty meters high."*

Do you think Dupuit's walls are a good analogy? In Dupuit's vision, how many "meters high" is consumer surplus?

*English translation of Jules Dupuit, "De la Mesure de l'Utilité des Travaux Publics," translated and reprinted in K. J. Arrow and T. Skitovsky, eds., *Readings in Welfare Economics* (Homewood, Ill.: Irwin, 1969).



Dupuit's Vision of Consumer Surplus

with an area of 1. The area is the extra amount that the consumer is getting because the market price is lower than what the consumer is willing to pay.

Consumer surplus for the entire market is the sum of the consumer surpluses of all individuals who have purchased goods in the market. In Figure 5.10, consumer surplus is the area between the market demand curve and the market price line.

Consumer surplus has many uses in economics. It is used to measure how well the market system works. We will show in Chapter 7 that the market system maximizes consumer surplus under certain circumstances. Consumer surplus can also be used to measure the gains to consumers from an innovation. For example, if a new production technique lowers the price of raisins, then the consumer surplus will increase: The area between the demand curve and the market price line increases. This increase is a measure of how much the new technique is worth to society.

Consumer surplus is also used to evaluate the benefits of government policies, such as building a new bridge or creating a new wilderness area. These policies will increase or decrease consumer surplus, and their value to society can be estimated using the concept of consumer surplus.

REVIEW

- Consumer surplus is the area between the demand curve and the market price line. It is a measure of how much the consumer gains from buying goods in the market.
- The consumer surplus for the market is the sum of the individual consumer surpluses.
- Consumer surplus is an important tool for measuring the performance of an economic system or for assessing the impact of alternative government policies.

Conclusion

This chapter is the first of three that looks at the individual behavior that underlies the economist's demand and supply curves. This chapter focused on consumers, Chapter 6 looks at firms, and Chapter 7 looks at the interaction of consumers and firms in markets. The payoffs in terms of understanding how and how well markets work will not be fully realized until we have completed all three chapters, but we have already derived a number of useful results.

We showed that the idea that people make purposeful choices with limited resources can be made operational with utility. We showed that utility maximization implies that a higher price reduces the quantity demanded.

By looking at consumers' willingness to pay, we showed that each consumer's demand curve for a good is given by the good's marginal benefit and that consumer surplus is the area under the demand curve and above the price.

KEY POINTS

1. The idea of utility maximization subject to a budget constraint implements the assumption that people make purposeful choices with limited resources.
2. Utility indicates the preferences people have for one activity compared with other activities.
3. The budget constraint shows the limit on how much a person can spend.
4. Economists assume—at least as an approximation—that people maximize their utility. To do so, people act as if they adjust their consumption to get to the highest level of utility.
5. Utility maximization implies that a higher price reduces the quantity demanded.
6. A person's demand curve can be derived from that person's utility.
7. Market demand curves are derived from individual demand curves.
8. Consumer surplus is the area between the demand curve and the market price line. Because a demand curve can be derived from willingness to pay and marginal benefits, consumer surplus is a measure of how much benefit a consumer gains from buying a product.

KEY TERMS

utility
budget constraint
utility maximization
income effect
substitution effect
marginal benefit
individual demand curve
market demand curve
consumer surplus

QUESTIONS FOR REVIEW

1. What is the relationship between utility and preferences?
2. Why don't the units by which utility is measured matter?
3. What is the relationship between utility maximization subject to a budget constraint and purposeful choice with limited resources?
4. Why does an increase in the price of a good reduce the number of combinations of goods a person can buy?
5. Why does a reduction in income lead to a reduction in the quantity demanded at each price?
6. Why are market demand curves usually smoother than individual demand curves?
7. What is the area below the demand curve and above the price?

PROBLEMS

1. Using the example of utility in Figure 5.2, find the quantity of each good the consumer will purchase in the cases shown in the table below.

Case	Budget	Price of Grapes	Price of Bananas
A	$7	$1	$1
B	$6	$2	$1
C	$8	$1	$2

2. Analyze the following data for Masa's utility from consumption of books and coffee.

Quantity of Books \ Quantity of Coffee	0	1	2	3	4
4	50	75	81	83	84
3	46	70	76	78	79
2	40	60	66	68	69
1	30	40	46	48	49
0	0	10	16	18	19

   a. Determine how much of each good Masa will consume if he has $20 and if the price of books is $10 and the price of coffee is $3.
   b. Suppose the price of coffee goes up to $5. How much coffee will Masa consume now? Why does the amount change?

3. Using the information from problem 2, multiply the utility received from books and coffee by 10. Will your answers to 2(a) and 2(b) change? Explain.

4. The following table shows Carl's willingness to pay for clothing.

Quantity of Clothing	Willingness to Pay
1	$35
2	$60
3	$80
4	$97
5	$112
6	$126

   a. Calculate Carl's marginal benefit from clothing.
   b. Draw Carl's individual demand curve for clothing.
   c. Suppose the price of one item of clothing is $17. How much would Carl consume, and what is his consumer surplus? Show your answer graphically as well as numerically.

5. The table below shows Andrew's willingness to pay for clothing.
   a. Draw Andrew's individual demand curve next to Carl's, and then add them together horizontally to derive a market demand curve.
   b. If the current market price of clothing is $20, how many items of clothing will each buy? Show this in the diagram.

Quantity of Clothing	Willingness to Pay
1	$70
2	$105
3	$130
4	$150
5	$167
6	$182

6. Consider the example of willingness to pay for X (raisins) in Table 5.1. Assume that the price is $.75 and that the person has $10 to spend. Compute the sum of the benefit (willingness to pay) and what the person has left over after paying for different amounts of raisins from 0 to 5 pounds. How many pounds of raisins will maximize the sum of the benefit plus the dollars left over? How does the answer compare to that using the price equals marginal benefit condition?

7. The following table shows Margaret's and Dennis's willingness to pay for cookies.

Quantity of Cookies	Margaret	Dennis
1	$7	$15
2	$13	$25
3	$18	$34
4	$21	$42
5	$23	$45

a. Calculate the marginal benefits for both people.
b. Derive Margaret's and Dennis's individual demand curves for cookies. Derive the market demand curve if only Margaret and Dennis are in the market.
c. Suppose that the price of cookies is $4.50. How many cookies will Margaret and Dennis buy? Calculate their consumer surplus. Draw a diagram like Figure 5.8 to show the area representing consumer surplus.
d. Show the consumer surplus for the whole market using the market demand curve. Draw a diagram like Figure 5.9.

8. Suppose that the willingness to pay for hearing aids by the economics professor mentioned in the introduction to this chapter was $60,000 for 1 pair, $60,400 for two pairs, $60,600 for three pairs, and $60,700 for four pairs. Draw the professor's demand curve for hearing aids. If the price of a pair of hearing aids is $500, how many pairs would the professor buy? What is the professor's consumer surplus? Show the consumer surplus in a diagram. Now suppose that a technological breakthrough reduces the price of hearing aids to $150 a pair. How many pairs will the professor buy now? What is the new consumer surplus?

Consumer Theory with Indifference Curves

Chapter 5 derives the demand curve from the assumption that consumers maximize utility subject to a budget constraint. Here we give a graphical illustration of that derivation.

Consider a single consumer deciding how much of two items to buy. Let one of the items be X and the other be Y. We first show that the consumer's budget constraint can be represented by a budget line, and then we show that the consumer's preferences can be represented by indifference curves.

The Budget Line

Suppose that the consumer has \$20 to spend on X and Y, and suppose that the price of X is \$2 per unit and the price of Y is \$4 per unit. How much of X and Y can the consumer buy? If the consumer spends all \$20 on Y, then 5 units of Y and no units of X are consumed. If the consumer buys 4 units of Y at \$4 per unit, then \$16 will be spent on Y and the remaining \$4 can be spent buying 2 units of X. These and several other amounts of X and Y that can be bought with \$20 are shown in the following table.

Units of Y	Units of X	Expenditures
5	0	$5 \times \$4 + 0 \times \$2 = \$20$
4	2	$4 \times \$4 + 2 \times \$2 = \$20$
3	4	$3 \times \$4 + 4 \times \$2 = \$20$
2	6	$2 \times \$4 + 6 \times \$2 = \$20$
1	8	$1 \times \$4 + 8 \times \$2 = \$20$
0	10	$0 \times \$4 + 10 \times \$2 = \$20$

These combinations represent the maximum amounts that can be purchased with \$20. Note that the amounts are inversely related; as more is spent on X, less must be spent on Y. This inverse relationship is shown graphically in Figure 5A.1. We put units of Y on the vertical axis and units of X on the horizontal axis, and then plot the pairs of points from the table. The points are then connected. The points trace a downward-sloping line starting at the upper left at $X = 0$ and $Y = 5$ and ending on the right with $X = 10$ and $Y = 0$. All the other combinations of X and Y in the table, such as $X = 4$ and $Y = 3$, are shown on the line. If it is possible to consume fractions of X and Y, then all the points on the line between the plotted points can also be purchased with the \$20.

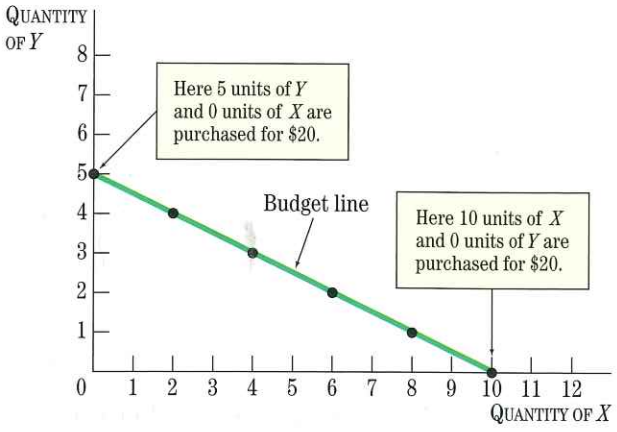


Figure 5A.1
Budget Line for a Consumer
The line shows how much a consumer with \$20 can consume of quantity X at a price of \$2 per unit and quantity Y at \$4 per unit. If \$20 is spent on Y and nothing on X, then 5 units of Y can be purchased, as shown on the vertical axis. If \$20 is spent on X and nothing on Y, then 10 units of X can be purchased. Other combinations are shown on the line.

(For example, 2.5 units of Y and 5 units of X would cost \$20: $2.5 \times \$4 + 5 \times \$2 = \$20$.) Because all these pairs of X and Y on this line can be purchased with a \$20 budget, we call it the **budget line.** The consumer is constrained to buy combinations of X and Y that are either on or below the budget line. Amounts of X and Y consumed below the budget line cost less than \$20. Points above the line require more than \$20 and are not feasible.

The budget line will shift out if the consumer has more to spend, as shown in Figure 5A.2. For example, if the consumer has \$24 rather than \$20, then the budget line will shift up by 1 unit because the extra \$4 permits the consumer to buy 1 more unit of Y. Alternatively, we could say that the budget line shifts to the right by 2 units in this case because the consumer can buy 2 more units of X with \$4 more.

The steepness of the budget line depends on the prices of X and Y. In particular, the slope of the budget line is equal to -1 times the ratio of the price of X to the price of Y. That is, slope $= -(P_X/P_Y)$, which is $-\frac{1}{2}$ in this example. Why is the slope determined by the price ratio? Recall that the slope is the change in Y divided by the change in X. Along the budget line, as X is increased by 1 unit, Y must fall by $\frac{1}{2}$ unit: Buying 1 more unit of X costs

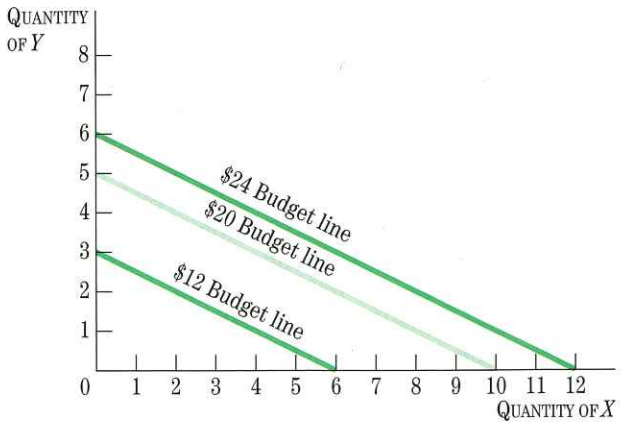


Figure 5A.2
Effect of a Change in Income on the Budget Line
If the consumer has more to spend, then the budget line is farther out. If the consumer has less to spend, then the budget line is farther in. Here a higher and a lower budget line are compared with the $20 budget line in Figure 5A.1.

$2 and requires buying ½ unit less of Y. Thus, the slope is $-½$.

In order to derive the demand curve for X, we need to find out what happens when the price of X changes. What happens to the budget line when the price of X increases from $2 to $4, for example? The budget line twists down, as shown in Figure 5A.3. The intuitive rationale for the twist is that the slope steepens to $-(P_X/P_Y) = -\$4/\$4 = -1$, and the position of $X = 0$ and $Y = 5$ on the vertical axis does not change, because 5 units of Y can still be purchased. You can show this by creating a new table with pairs of X and Y that can be purchased with $20 at the new price and then plotting the points.

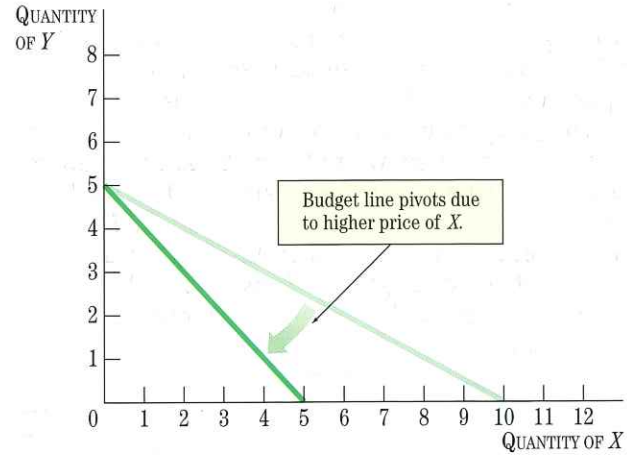


Figure 5A.3
Effect of a Higher Price of X on the Budget Line
The budget line pivots if the price of X changes. Here the price of X rises from $2 to $4 and the budget line twists down.

To summarize, we have shown how a budget line represents the budget constraint for the consumer; now we show how to represent the consumer's preferences.

The Indifference Curve

Utility is an indicator of how a consumer prefers one combination of items in comparison with another. If the level of utility is the same for two combinations of X and Y, then the consumer is *indifferent* between the two combinations. Suppose that the utility is the same for the combinations of X and Y that appear below.

Units of Y	Units of X
6	1
4	2
2	6
1	12

The consumer is indifferent among these combinations. Observe that these amounts are inversely related. As consumption of Y declines, the consumer must be compensated with more X if the level of utility is not to decline.

We can plot these different amounts on the same type of graph we used for the budget line, as shown in Figure 5A.4. The consumer is indifferent among all four points. We have connected the points with a curve to

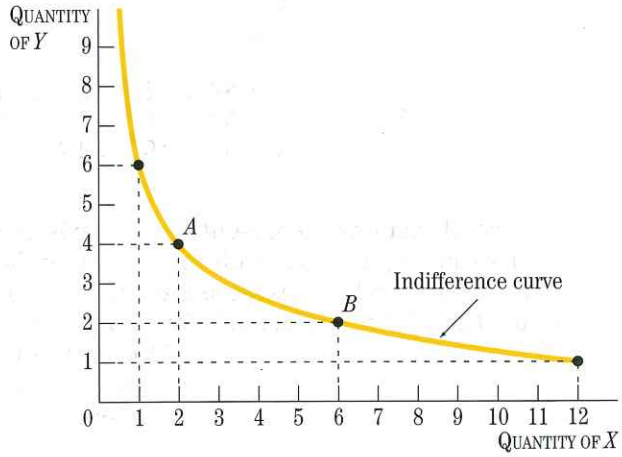


Figure 5A.4
An Indifference Curve for a Consumer
The consumer is indifferent between A and B or any other point on an indifference curve. For example, the consumer is indifferent between consuming 4 units of Y and 2 of X or 2 units of Y and 6 of X.

represent other combinations of X and Y about which the consumer is indifferent. The curve is called an **indifference curve** because the consumer is indifferent among all points on the curve. The indifference curve slopes downward from left to right.

The slope of the indifference curve can be found from the marginal utility, X and Y. **Marginal utility** is *the increase in utility from consuming an additional unit of a good.* For example, look back at Figure 5.2 and consider the marginal utility of increasing consumption of grapes by 1 additional pound, from 3 pounds to 4 pounds. You will see that utility increases by 2 when grape consumption increases from 3 pounds to 4 pounds. Thus, the marginal utility of grapes is 2 at the amount of consumption. Let MU_X be the marginal utility of X and let MU_Y be the marginal utility of Y.

The slope of the indifference curve is equal to negative 1 times the ratio of the marginal utility of X to the marginal utility of Y; that is, slope $= -(MU_X/MU_Y)$. The reason is that utility is the same for all points on an indifference curve. In other words, the decline in utility as X falls $(-MU_X \times \Delta X)$ must equal the increase in utility as Y rises $(MU_Y \times \Delta Y)$. Thus, $(MU_X \times \Delta X) = -(-MU_Y \times \Delta Y)$, or $-MU_X/MU_Y = \Delta Y/\Delta X$, which is the slope of the indifference curve.

Note that the indifference curve is bowed in toward the origin. That is, the indifference curve is steep when a small amount of X is consumed and flat when a large amount of X is consumed. This curvature is due to the declining marginal rate of substitution. When the consumer is consuming only a little bit of X, a large amount of Y is required as compensation for a reduction in X. As X increases, less of Y is required as compensation.

The ratio of marginal utilities MU_X/MU_Y is called the **marginal rate of substitution;** it gives the number of units of one good (Y) for which the consumer is willing to trade one unit of the other good (X) and have the same amount of utility—or be indifferent. For example, if the marginal rate of substitution is 4, then the consumer is willing to trade 4 units of Y for 1 unit of X with utility remaining the same.

We can represent higher levels of utility or more preferred combinations of X and Y by higher indifference curves, as shown in Figure 5A.5. Any point on a higher indifference curve is preferred to any point on a lower indifference curve.

Getting to the Highest Indifference Curve Given the Budget Line

Now we can combine the budget line and the indifference curves on the same diagram to illustrate the model of consumer behavior. Utility maximization subject to the budget constraint means getting to the highest possible indifference curve without going above the budget line. The process is shown in Figure 5A.6. The budget line

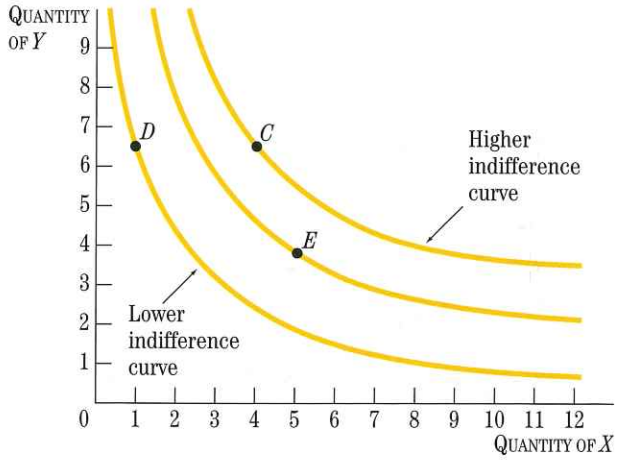


Figure 5A.5
Higher and Lower Indifference Curves
Amounts of X and Y on indifference curves that are higher are preferred to amounts on indifference curves that are lower. Of the three combinations C, D, and E, the combination at D is the least preferred and the combination at C is the most preferred.

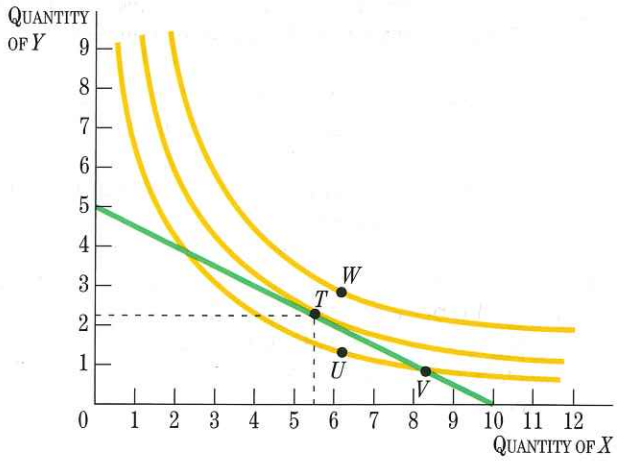


Figure 5A.6
The Best Choice for the Consumer
When the budget line is tangent to the indifference curve, the consumer cannot do any better. The point of tangency is at point T. Compare this with the other points. Point U is not the best point because it is inside the budget line. Point V is not the best point because there are other points on the budget line that are preferred. Point W is preferred to point T, but it is not feasible.

from Figure 5A.1 and the indifference curves from Figure 5A.5 are shown in the diagram. The consumer cannot go beyond the budget line, and any point inside the budget line is inferior to points on the budget line. Thus the combination of X and Y with the highest utility must be on the budget line. The highest indifference curve with points on the budget line is the one that just touches—is tangent to—the budget line. This occurs at point T in Figure 5A.6. The **tangency point** is the highest level of utility the consumer can achieve subject to the budget constraint. It is the combination of X and Y that the consumer chooses. Figure 5A.6 shows that, in this example, the consumer buys 2¼ units of Y and 5½ units of X.

The Utility-Maximizing Rule

Observe that at the tangency point, the slope of the budget line is equal to the slope of the indifference curve. That is, $P_X/P_Y = MU_X/MU_Y$. In other words, the ratio of the price of two goods equals the ratio of the marginal utility of the two goods as long as the consumer is maximizing utility. This equality between the price ratio and the ratio of the marginal utilities, or the marginal rate of substitution, is called the *utility-maximizing rule.*

Effect of a Price Change on the Quantity Demanded

Now suppose that the price of X increases; then the budget line twists down, as shown in the lower panel of Figure 5A.7. With the new budget line, the old consumer choice of 2¼ units of X and 5½ units of Y is no longer feasible: Point T is outside the new budget line. The highest level of utility the consumer can now achieve is at point S in the lower panel of Figure 5A.7. At point S, the quantity of X has declined. Thus, a higher price of X has reduced the quantity of X demanded.

In the top graph in Figure 5A.7, we show the relationship between the price of X and the quantity of X demanded. The price of X is put on the vertical axis, and the quantity of X demanded is put on the horizontal axis. The lower quantity demanded at the higher price shows the negative slope of the demand curve.

Effect of an Income Change on Demand

We can also examine what happens when the consumer's income changes but the price remains constant. This is illustrated in Figure 5A.8, where income declines. The lower income leads to less consumption of both X and Y. In this case, both X and Y are normal goods because consumption goes down when income goes down. If the consumption of X increased as the budget curve shifted in, then X would be an inferior good.

Graphical Illustration of Income Effect and Substitution Effect

The effect of a change in the price on the quantity demanded can be divided into an income effect and a substitution effect. These two effects can be represented graphically as shown in Figure 5A.9.

As in Figure 5A.7, there is a twist in the budget line due to the higher price of X. But now we draw in another budget line—the dashed line in Figure 5A.8—corresponding to a lower level of income due to the higher price of X, but without twisting the line. This budget line has the same slope as the original line. Thus the dashed budget line shows the reduction in income due to the price increase of X but keeps the relative price of X to Y the same as it was before the price of X increased. The

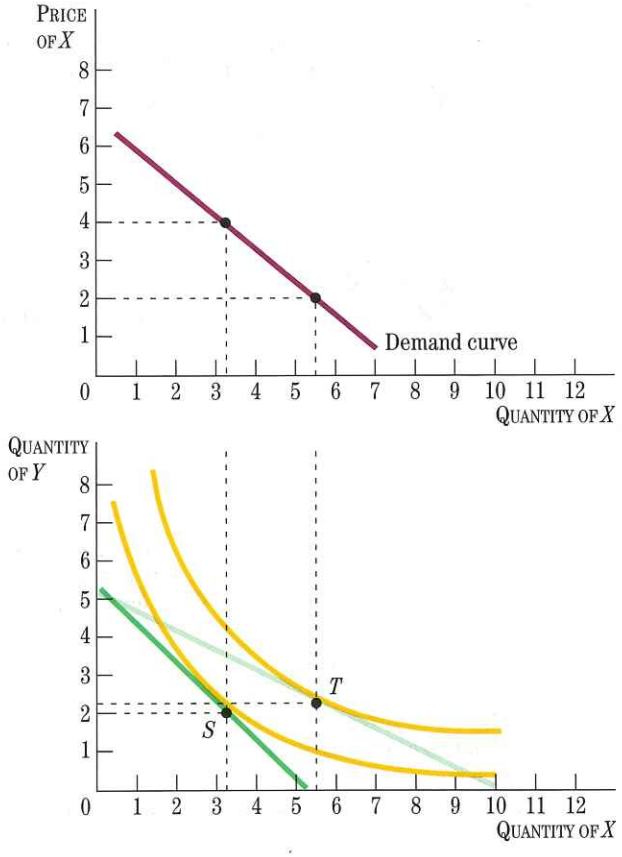


Figure 5A.7
An Increase in the Price of *X*
If the price of X rises, the budget line pivots down and the consumer's choice changes from point T to point S in the lower panel. The quantity of X consumed goes down. The price of X and the quantity of X are plotted in the top panel, showing the negative relationship between price and quantity demanded.

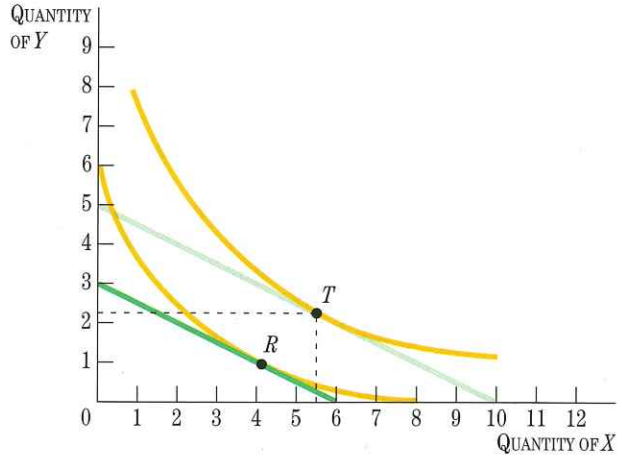


Figure 5A.8
A Decrease in Income
If the consumer's income falls, there is a new point at which utility is maximized: The consumer moves from point *T* to point *R*. In this case, consumption of both *X* and *Y* declines. Neither good is an inferior good in this example.

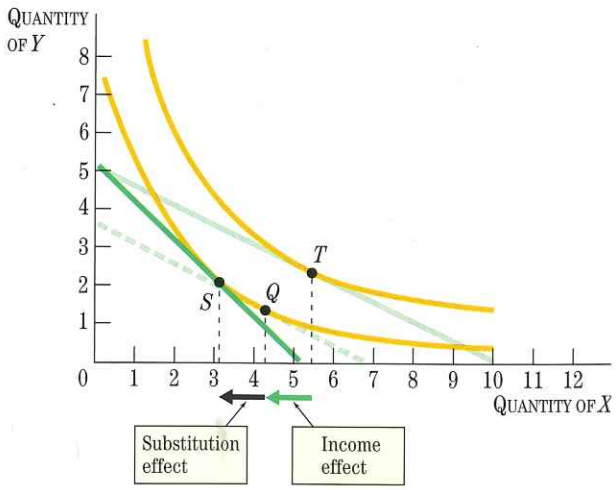


Figure 5A.9
Illustration of Income Effect and Substitution Effect of a Price Change
The dashed budget line has the same slope as the original line and leads to the income effect. The rest of the decline in the quantity of *X* is the substitution effect.

dashed line can be used to find the income effect without the substitution effect because it has the same slope as the original budget line.

Observe that the tangency of this dashed budget line with the indifference curve at *Q* gives a lower level of consumption of *X* compared with the original point *T*. The decline of consumption from *T* to *Q* is the income effect. The remaining decline from *Q* to *S* is the substitution effect.

Key Points

1. The budget line represents the consumer's budget constraint in a diagram with the quantity consumed of each of two goods on the axes.
2. The budget line is downward-sloping, with the slope equal to negative 1 times the ratio of the price of the good on the horizontal axis to the price of the good on the vertical axis.
3. A higher price of the good on the horizontal axis twists the budget line down.
4. An indifference curve shows the combinations of goods among which the consumer is indifferent.
5. Combinations of goods on higher indifference curves are preferred to combinations of goods on lower indifference curves.
6. The model of consumer behavior assumes that the consumer tries to get to the highest possible indifference curve without going beyond the budget line.
7. The consumer chooses the combination at the tangency of the budget line and the indifference curve.
8. A higher price of a good lowers the quantity demanded, according to the indifference curve and budget line diagram.

Key Terms and Definitions

budget line: a line showing the maximum combinations of two goods that it is possible for a consumer to buy, given a budget constraint and the market prices of the two goods.

indifference curve: a curve showing the combinations of two goods that leave the consumer with the same level of utility.

marginal utility: the increase in utility when consumption of a good increases by one unit.

marginal rate of substitution: the amount of one good for which the consumer is willing to trade one unit of another good and still have the same utility.

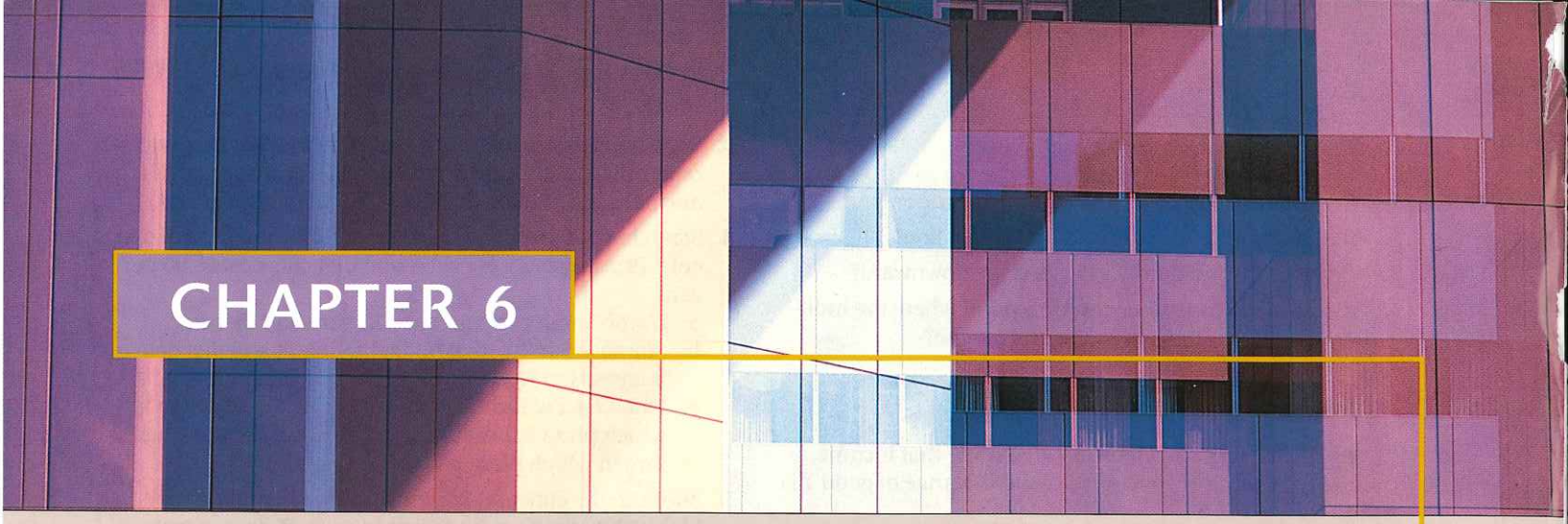
tangency point: the only point in common for two curves; the point where the two curves just touch.

Questions for Review

1. Why does the budget line slope downward?
2. What determines the slope of the budget line?
3. Why does the indifference curve slope downward?
4. Why does the consumer choose a point where the indifference curve is tangent to the budget line?

Problems

1. Draw a diagram like Figure 5A.7. Assume that income is \$20, the price of good X is \$2, and the price of good Y is \$4.
   a. Show what happens to the budget line if the price of X falls from \$2 to \$1.60. What is the maximum amount of X that can be purchased?
   b. Illustrate the new point of consumer choice (by drawing the appropriate indifference curve).
   c. Show the demand curve for the consumer.
2. Darnell has \$30 to spend on either muffins, which cost \$3 each, or cartons of orange juice, which cost \$1.50 each.
   a. Graph Darnell's budget line for muffins and orange juice. What is the maximum quantity of orange juice that Darnell can buy with \$30?
   b. Suppose the price of orange juice increases to \$2 per carton. Show the change in the budget line.
   c. Draw indifference curves to show the change in consumption of orange juice, assuming that orange juice is a normal good.
3. Suppose the consumer in Figure 5A.7 has \$24 to spend rather than \$20. Draw two diagrams—one in which X is a normal good and one in which X is an inferior good.
4. Sarah has \$20 to spend on slices of pizza and cans of diet cola. Pizza costs \$1 per slice, and diet cola costs \$.50 per can.
   a. Graph Sarah's budget line for pizza and diet cola.
   b. Suppose Sarah's total budget for pizza and diet cola increases to \$25. How does her budget line shift?
   c. Draw a set of indifference curves for the situation in which pizza is a normal good, and one for the situation in which pizza is an inferior good.
5. Suppose the consumer is consuming at a point on the budget line that is not tangent to an indifference curve. Explain why utility is not being maximized at this point.



CHAPTER 6

The Supply Curve and the Behavior of Firms

Americans love their bumper stickers. They buy tens of thousands of them every year to proclaim their loyalty to movements, political candidates, schools, or a particular way of life. But some people who would like to express their views on their car might hesitate before applying a sticker that could be hard to remove later on. So when yellow-ribbon car magnets proclaiming support for U.S. troops in Iraq showed up in stores in 2003, many people eagerly bought them, paying far more for the magnets (around $5 each) than the pennies it cost to produce them—creating, in economic terminology, a *producer surplus*. Dwain Gullion, who began the whole craze with the casual production of 1,000 magnets he intended to sell at his Christian bookstore, had to gear up production fast. By the summer of 2004, Gullion's new company, Magnet America, was shipping more than 100,000 magnets a week and employing more than 100 contract workers. Other firms quickly got into the act, and soon Gullion had some hard decisions to make about his business.

In this chapter, we look at the behavior of firms such as Magnet America. The behavior of firms can be described by supply curves like the one shown in Figure 6.1. The supply curve tells us how much all the firms in the market—not just a single firm like Magnet America—would produce at each price.

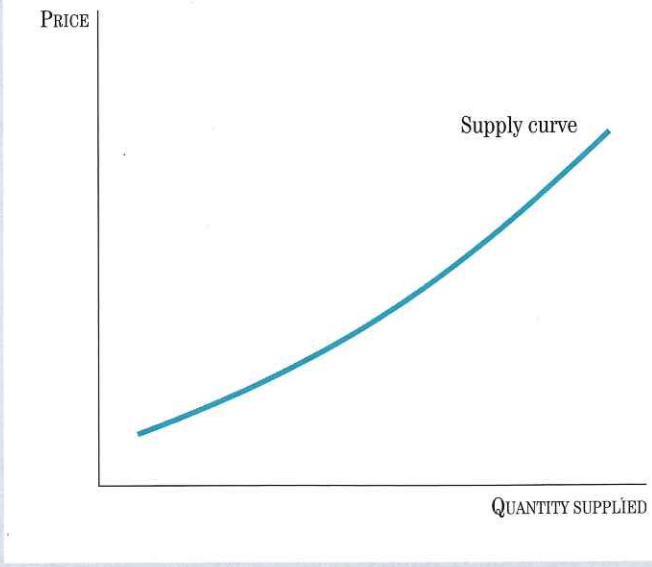
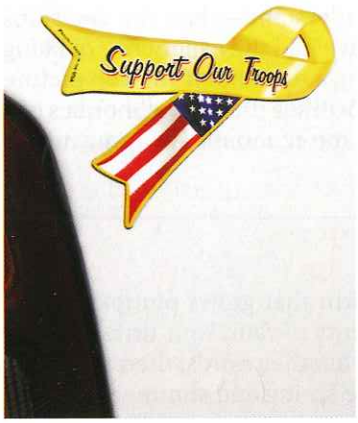


Figure 6.1
A Typical Supply Curve for a Market
Supply curves typically slope upward. In this chapter, we look at the factors that motivate firms in the market to increase the quantity supplied as the price rises.

As with our study of the consumers who underlie the demand curve in Chapter 5, there are two important reasons to study the firms that underlie the supply curve. First, we want to show why the supply curve has the slope and position it does. Such information enables us to assess how a change in technology or trends or a new government policy affects the supply curve. Second, we want to show how and why a supply curve can be used to measure the "producer surplus" of firms.

This chapter on supply parallels Chapter 5, on demand, and we will exploit that parallel to make our analysis easier. Recall from Chapter 1 that the central idea of economics is that *people make purposeful choices with scarce resources and interact with other people when they make these choices.* In Chapter 5, we looked at consumers making purposeful choices with scarce resources by assuming that they maximize utility subject to a budget constraint. In this chapter, we look at firms making purposeful choices with scarce resources by assuming that they maximize profits subject to a production function (see Figure 6.2). In Chapter 7, we consider the interaction of consumers and firms in markets as they make these choices.



Basic Economic Principle	When Applied to the Behavior of Consumers	When Applied to the Behavior of Firms
People...	*Consumers...*	*Firms...*
make purposeful choices...	maximize utility...	maximize profits...
with scarce resources.	subject to a budget constraint relating expenditure to income.	subject to a production function relating output to input.

Figure 6.2
Applying the Central Idea of Economics
People make purposeful choices with scarce resources and interact with other people when they make these choices. In Chapter 5, the people were the consumers. In this chapter, they are the firms. In Chapter 7, the consumers and the firms interact with each other.

The new terms introduced in this chapter have analogies with terms introduced in Chapter 5. For example, we will explicitly define producer surplus and compare it to consumer surplus. We will derive a supply curve for an individual firm, which will tell us the quantity of a good the firm would supply at different prices. The supply curve for an individual firm is analogous to the demand curve for an individual consumer, which tells us the quantity of a good the individual consumer demands at different prices.

Definition of a Firm

firm: an organization that produces goods or services.

We start by looking at the behavior of a single firm. A **firm**, by definition, is an organization that produces goods or services. Just as no two consumers are exactly alike, no two firms are exactly alike. A firm can be a small family farm in the country or a grocery store in the city. Bakeries, restaurants, auto dealers, and bicycle shops are all examples of firms that are usually relatively small. Other firms—such as General Motors—are very large, producing many different products in large volume.

The terms *firm, company,* and *business* are used interchangeably. A firm may include several *establishments*, which are separate physical locations, such as an office, a factory, or a store, where work is done. For example, the U.S.-based grocery chain Kroger is a firm with over 2,000 establishments—more than 1,000 supermarkets and more than 1,000 convenience stores, including 208 Kwik-Starts and 105 Mini-Marts. Of course, many small firms have only one establishment.

In the United States, about 80 percent of all firms are *sole proprietorships*, with one owner, or *partnerships*, with only a few owners, who usually manage the firm. Most of these are very small when compared with corporations like General Motors. A *corporation* is unlike a sole proprietorship or partnership in that the managers are usually somewhat removed from the owners. For example, most people who own shares of General Motors never even meet the managers of the firm. This separation of managers and owners means that the managers must be given an incentive to keep the owners' interests in mind. A common incentive is to have managers share in the profits of the firm.

You might expect that the decisions made by the managers of a firm are more complicated—and consequently more difficult to understand—than the decisions made by consumers. Of course, many more people have had the experience of being a consumer than have had the experience of managing a firm. But if you can picture yourself as the owner/manager of your own firm, you will see that the economics of a firm's decision about how much to sell is analogous to the economics of a consumer's decision about how much to buy.

Your Own Firm: A Pumpkin Patch

Imagine that you are the owner and manager of a firm that grows pumpkins on a pumpkin patch; the patch has good soil and gets plenty of rain. Your firm is one of many specializing in growing and selling pumpkins—in other words, there are many other firms with which you must compete. During the spring and summer you grow the pumpkins, and in the fall you sell them. As owner and manager of the firm, you must pay rent at the start of each growing season to the landlord who owns the

pumpkin patch. During the season, you hire workers to tend the patch. The more workers tending the pumpkins, the more pumpkins you can grow on the patch.

Your firm is typical of many small firms and has features that apply to larger firms as well. Your firm is one with a single product (pumpkins) and two factors of production—land (the patch) and labor (the workers). One of the factors of production, land, cannot be changed during the season because the rent was paid in advance; this makes land a *fixed factor*. The other factor, labor, can be varied during the season, because you can choose to hire more or fewer workers; this makes labor a *variable factor*.

Your Firm as a Price-Taker in a Competitive Market

Keep your firm in mind as you go through this chapter. It will be referred to often. Our aim is to determine the supply curve for a firm like your firm. *A supply curve for a single firm tells us the quantity of a good that that firm will produce at different prices.* To find the supply curve, we imagine that the firm looks at the price of the good it is selling and then decides how much to produce. For example, a baker considers the price of a loaf of bread prevailing in the market when deciding how many loaves of bread to produce. Because the firm takes the price in the market as a given when it makes a decision about production, we say that the firm is a **price-taker.** We assume that your pumpkin firm is a price-taker. This means that you decide how much to produce and sell after looking at the price of pumpkins in the market.

price-taker: any firm that takes the market price as given; this firm cannot affect the market price because the market is competitive.

This description of a firm as a price-taker may seem odd to you. After all, if the firm doesn't set the price, then who does? Of course, in some sense, each firm does. If you go to a bakery for a loaf of bread, a price tag states the price of the loaf, so the baker is clearly determining the price. But this is not the way economists look at it; there is a subtlety here in the way economists describe the market. When there are many bakers selling bread, in an important sense, the individual bakers do not have the ability to affect the price by much. If one baker charges $3 for a loaf of bread, and all the other bakers in the community charge $1.50 for the same loaf, no one will buy bread from the first baker. People will not even go to the store if they know the price is that high. They will go to other bakeries, where bread sells for $1.50 a loaf. Although in principle an individual firm has the ability to set any price it wants, in reality a firm cannot charge a price far from the price that prevails in the market without soon losing all its customers.

A market in which a single firm cannot affect the market price is called a **competitive market.** The market for fresh bread, with many bakeries in any reasonably sized community, is competitive. Because many firms are producing pumpkins along with your pumpkin firm, the pumpkin market is also competitive. A competitive market requires that there be at least several firms competing with one another. Exactly how many firms are required to make a market competitive is difficult to say without studying the market carefully—as we do in later chapters. If a market is competitive, so that firms are price-takers, then we can derive a supply curve by asking, "How much bread would the baker produce if the price of bread were $1 a loaf? $1.50 a loaf? $2 a loaf?" and so on. Or, in the case of your pumpkin firm, "How many pumpkins would the pumpkin firm produce if the price of pumpkins were $35 a crate? $70 a crate?" and so on.

competitive market: a market in which no firm has the power to affect the market price of a good.

Other Types of Markets

Not all markets are as competitive as the fresh bread market or the pumpkin market, and part of our job later in the book is to study these markets. The exact opposite of a competitive market is where there is only one firm, in which case the firm is called a

monopoly. Strictly speaking, a monopoly does not have a supply curve because the monopoly does not take the price as given. Instead, the monopoly can dictate the price (it is a *price-maker*). The question "How much does the monopoly produce at a given price?" has no meaning because the monopoly need not take the price as given. We consider monopolies in Chapter 10. For now we focus on the price-taking firms in a competitive market.

This subtlety about firms taking prices as given does not seem to arise in the case of the consumer. In deriving the demand curve in Chapter 5, we assumed that the individual consumer could not affect the price. This seems natural because we do not usually see buyers setting the price for bread or other commodities. As long as there are at least several buyers and several sellers in the market, we can assume that the price is taken as given by both buyers and sellers. In Chapter 7, when we study the interaction of buyers and sellers in markets, we will show how the market price is determined.

REVIEW

- There are a great variety of sizes and types of firms.
- A market is competitive if no single firm can affect the market price.
- In a competitive market with many firms, each firm is a price-taker.
- The supply curve of a firm describes how the quantity produced depends on the price.

The Firm's Profits

profits: total revenue received from selling the product minus the total costs of producing the product.

Profits for any firm—a bakery producing bread or a farm producing pumpkins—are defined as the *total revenue* received from selling the product minus the *total costs* of producing the product. That is,

$$\text{Profits} = \text{total revenue} - \text{total costs}$$

When profits are negative—total revenue is less than total costs—the firm runs a *loss*. When profits are zero—total revenue is equal to total costs—the firm is *breaking even*.

We assume that the firm *maximizes* profits. That is, the firm decides on the quantity of production that will make profits as high as possible. To see how this is done, we must examine how profits depend on the quantity produced. To do this, we must consider how total revenue and total costs—the two determinants of profits—depend on the quantity produced. We consider first total revenue and then total costs.

Total Revenue

total revenue: the price per unit times the quantity the firm sells.

Total revenue is the total number of dollars the firm receives from people who buy its product. Total revenue can be computed by multiplying the price of each unit sold by the quantity sold. That is,

$$\begin{aligned}\text{Total revenue} &= \text{price} \times \text{quantity} \\ &= P \times Q\end{aligned}$$

Table 6.1
Total Revenue from Pumpkin Production at Three Prices

Quantity Produced (crates)	Total Revenue		
	Price = $35/crate	*Price = $70/crate*	*Price = $100/crate*
0	0	0	0
1	35	70	100
2	70	140	200
3	105	210	300
4	140	280	400
5	175	350	500

where we use the letter P to stand for price and Q to stand for quantity. Because we are looking at an individual firm and a particular product, P is the price of the particular product the individual firm is selling, and Q is the number of items the firm sells. There are a variety of ways to measure the quantity sold: numbers of crates of pumpkins, slices of pizza, loaves of bread, quarts of milk, kilowatts of electricity, and so on. In the United States, pounds and tons are usually used to measure items like coal, grapes, wheat, and sugar, but kilos would do as well.

Note that total revenue depends both on the price of the item being sold and on the number of items sold. For example, if the price of bread is $1.50 per loaf and 100 loaves are sold, then total revenue is $150. If the price rises to $2 a loaf and 200 loaves are sold, then total revenue is $400. The more items sold at a given price, the higher total revenue is. Thus, the firm can increase total revenue by producing and selling more goods.

Table 6.1 shows how total revenue increases with the quantity produced for your firm producing pumpkins. Each row of the table shows the total revenue the firm receives from selling varying amounts of pumpkins. Each column showing total revenue corresponds to a different price: $35 per crate, $70 per crate, and $100 per crate. For example, when the firm can get $70 per crate, it receives $280 for selling 4 crates.

Production and Costs

total costs: the sum of variable costs and fixed costs.

Now that we have seen how total revenue depends on the quantity produced, let's examine how total costs depend on the quantity produced. **Total costs** are what the firm has to incur in order to produce the product. For your pumpkin firm, total costs include the workers' salaries and the rent on the land. To see how total costs depend on the quantity produced, we must look at what happens to the quantity of labor and land used by the firm when the quantity produced increases or decreases.

Reminder: In Chapter 1, we saw that costs include *opportunity costs*. Thus, total costs for your pumpkin firm would include the opportunity cost of any time you spent operating the firm rather than doing something else, like studying for an exam. To emphasize that opportunity costs are included in total costs when computing profits, economists sometimes use the term *economic profits* rather than simply profits.

■ **The Time Period.** We look at the firm's production decisions over a short period of time—such as one growing season—rather than over a long period of time—such as several growing seasons. Because we focus on the short run, we assume that only the labor input to production can be varied. Our analysis of the firm in this chapter is called a *short-run* analysis because the time is too short to change the other factors of production, such as land; only labor can be changed. We make this assumption simply because it is easier to examine the firm's decisions when only one factor of production can be changed. It is a simplifying assumption that we will modify. In Chapter 8 we take up the *long run*, in which other factors of production—such as the size of the pumpkin patch—can change as well as labor.

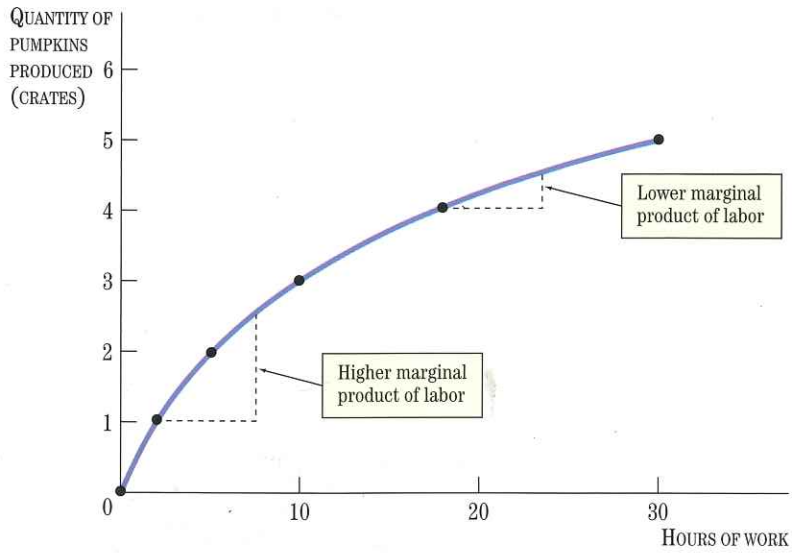


Figure 6.3
A Production Function Relating Output to Labor Input
As more workers are employed, production increases. But the increase in production added by each additional hour of work declines as more workers are hired because the land the workers have to work with does not increase. Thus, there is a decreasing marginal product of labor, or diminishing returns to labor.

production function: a relationship that shows the quantity of output for any given amount of input.

marginal product of labor: the change in production due to a one-unit increase in labor input.

diminishing returns to labor: a situation in which the increase in output due to a unit increase in labor declines with increasing labor input; a decreasing marginal product of labor.

■ **The Production Function.** Figure 6.3 plots the relationship between pumpkin production and labor input. The number of hours of work is on the horizontal axis, and the quantity produced is on the vertical axis. Each point in Figure 6.3 shows the number of hours of work and the quantity of pumpkins produced: To produce 3 crates requires 10 hours of work; to produce 5 crates requires 30 hours of work. Clearly, more pumpkin production requires more labor input. The graph in Figure 6.3 is called the firm's **production function** because it tells us how much is produced for each amount of labor input. This production function is for a given size of pumpkin patch.

The **marginal product of labor** is defined as the increase in production that comes from an additional unit of labor. Figure 6.3 shows that the marginal product of labor *declines* as labor input increases. Because of the curvature of the production function, the same increase in hours of work leads to a smaller increase in production when labor input is large than when labor input is small.

Another term for the phenomenon of declining marginal product of labor is **diminishing returns to labor.** In your pumpkin firm, diminishing returns to labor occur as additional workers are employed. As more and more workers are employed on a given amount of land, each additional worker adds less and less additional output. Diminishing returns is a general phenomenon that occurs when some inputs to production—such as land or machines—are fixed. Because the size of your pumpkin patch is fixed, additional workers must eventually add less to production. Otherwise a single plot of land could produce all the world's pumpkins by employing huge numbers of workers. Thus, the return to each additional worker declines. Diminishing returns to labor occur in nonagricultural examples as well. Employing more and more workers at a single McDonald's will increase the amount of fast food produced by less and less.

■ **Costs.** Table 6.2 shows how total costs depend on the quantity of pumpkins produced at your pumpkin firm. The first column shows the quantity of pumpkins

Table 6.2
Example of Costs for a Single Firm

Quantity Produced (crates)	Hours of Labor Input	Variable Costs at $10 Wage (dollars)	Fixed Costs (dollars)	Total Costs (dollars)
0	0	0	50	50
1	2	20	50	70
2	5	50	50	100
3	10	100	50	150
4	18	180	50	230
5	30	300	50	350

produced. The second column shows the labor input required to produce that quantity of pumpkins, using the production function from Figure 6.3. The other columns show how total costs are determined.

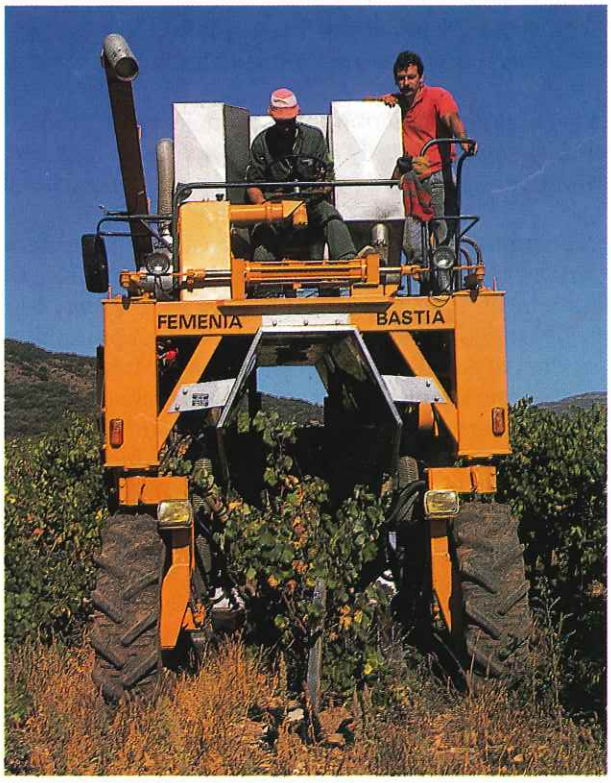
The first row shows how much it costs if you decide to produce zero pumpkins. We assume that you have to pay $50 up front for rent on the patch even if you decide to produce no pumpkins. These payments are considered *fixed costs* because they must be paid no matter how many pumpkins are produced. By definition, **fixed costs** are the part of total costs that do not depend on how much is produced.

fixed costs: costs of production that do not depend on the quantity of production.

The next row of Table 6.2 shows the costs of producing 1 crate of pumpkins. The additional costs of producing 1 crate compared to zero crates are $20. That is payment for 2 hours of work at $10 per hour. The $20 in payments are **variable costs** because they vary according to how much is produced. These costs are variable because more workers are hired as more is produced. Variable costs and fixed costs together constitute all the costs of producing the product and must be subtracted from revenue to get profits. Hence, *the sum of fixed costs and variable costs equals total costs*, as shown in the last column of Table 6.2.

variable costs: costs of production that vary with the quantity of production.

The third row of Table 6.2 shows the costs of producing 2 crates of pumpkins. Clearly 2 crates are more costly to produce than 1 crate because more workers are needed. Variable costs, with 5 hours of work at $10 per hour, rise to $50. As more pumpkins are harvested, more workers must be hired, and the total costs increase. The remaining rows of Table 6.2 show what happens to costs as the quantity produced increases further.



Diminishing Returns to Labor
Adding the second worker to this machine in a French vineyard increased the quantity of grapes produced by much less than the first worker did. Adding a third worker to the machine would increase the quantity of grapes produced by even less than adding the second worker did. Thus, this is an example of a decreasing marginal product of labor, or diminishing returns to labor.

Table 6.3
Total Costs and Marginal Cost

Quantity Produced (crates)	Total Costs (dollars)	Marginal Cost (dollars)
0	50	—
1	70	20
2	100	30
3	150	50
4	230	80
5	350	120

The connecting lines emphasize how marginal cost is the change in total costs as the quantity produced increases by one unit.

marginal cost: the change in total costs due to a one-unit change in quantity produced.

Marginal cost is defined as the increase in total costs associated with an additional unit of production. Table 6.3 shows how marginal cost is calculated for the example in Table 6.2. For example, the marginal cost of increasing production from 1 crate to 2 crates is \$30 (\$100 − \$70 = \$30), and the marginal cost of increasing production from 2 crates to 3 crates is \$50 (\$150 − \$100 = \$50).

Notice how marginal cost increases as production increases. Marginal cost is greater when we go from 2 to 3 crates (\$50) than when we go from 1 to 2 crates (\$30). The pattern of *increasing marginal cost* is apparent throughout the range of production in Table 6.3.

Observe that *increasing marginal cost is due to the diminishing marginal product of labor:* The marginal cost of going from 2 crates to 3 crates is greater than that of going from 1 crate to 2 crates because more worker hours are required to raise production from 2 crates to 3 crates than are required to raise production from 1 crate to 2 crates.

Increasing marginal cost is a general phenomenon that occurs in many production processes. It is essential for deriving the supply curve. In fact, as we will see, increasing marginal cost is the whole reason that the supply curve for an individual firm slopes upward.

There are exceptions to the principle of increasing marginal cost. One important exception is that marginal cost need not be increasing over the entire range of production. For example, there might be a decrease in marginal cost at very low levels of production. If a team of at least two workers is needed to harvest pumpkins, for example, then the marginal product of a second worker might be greater than the marginal product of a first worker. One worker might add very little, whereas the second might add a lot. But diminishing returns to labor and increasing marginal cost eventually set in as more workers are hired and more pumpkins are produced.

This chapter assumes that marginal cost increases over the whole range of production. This is a common assumption used by economists and is a good approximation except for very low levels of production. In Chapter 8, we will see what happens when marginal cost declines at low levels of production.

■ **Graphical Representation of Total Costs and Marginal Cost.** A better understanding of how a firm's total costs depend on production can be obtained by representing the total costs graphically. Figure 6.4 plots the pairs of numbers for total costs and quantity produced from the first two columns of Table 6.3. Dollars are on the vertical axis, and the quantity of pumpkins produced is on the horizontal axis. Note how the total costs curve bends up: As marginal cost increases, the curve gets

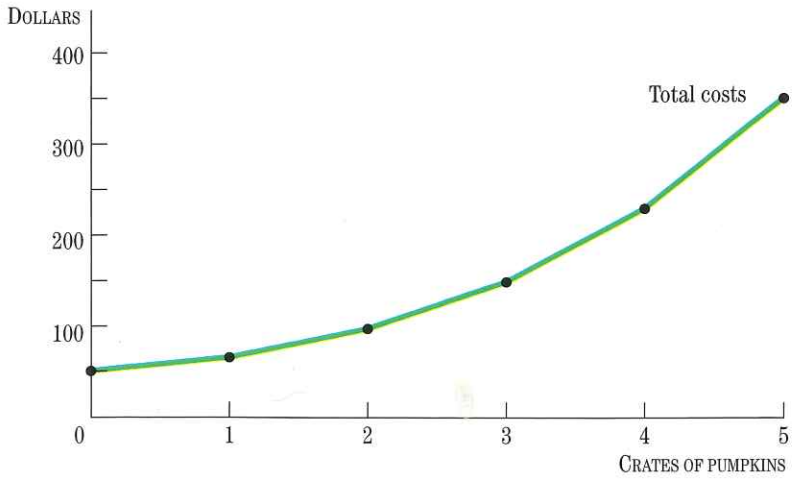


Figure 6.4
Total Costs
In order to produce goods, a firm incurs costs. For example, more workers must be paid to produce more goods. As more goods are produced, the firm's total costs rise, as shown here. At higher levels of production, costs increase by larger amounts for each additional item produced.

steeper, or the slope increases. The marginal cost is the slope of the total costs curve. The increasing slope is a visual way to show the increasing marginal cost.

Figure 6.5 shows the relationship between marginal cost and number of crates of pumpkins produced. The points in Figure 6.5 correspond to the pairs of numbers in the first and third columns of Table 6.3. Note that the marginal cost curve slopes upward, illustrating how marginal cost increases. The firm's objective to maximize

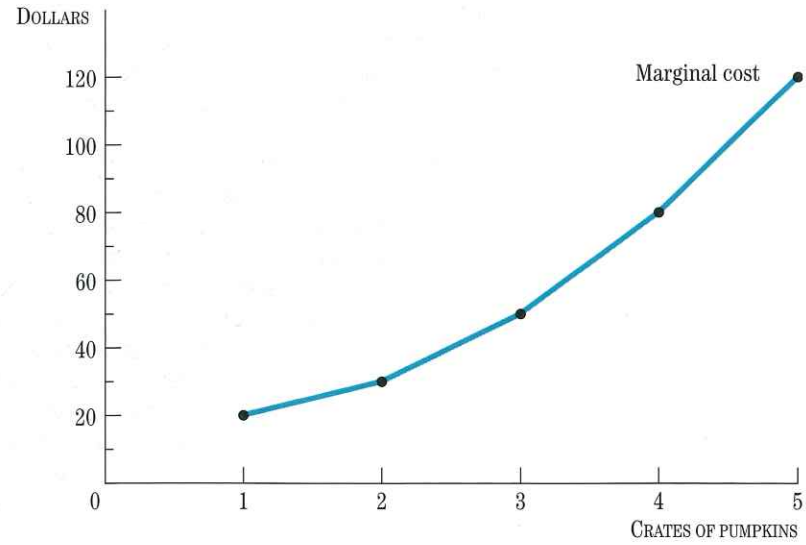


Figure 6.5
Marginal Cost
The change in total costs as more units of the good are produced is called marginal cost. Marginal cost increases as more units are produced, as illustrated here.

profits, coupled with the shape of the marginal cost curve, determines the upward-sloping supply curve, as we explain in the next section.

REVIEW

- Profits are defined as the difference between total revenue and total costs.
- Profits depend on the quantity produced because both total revenue and total costs depend on the quantity produced.
- Total revenue is defined as the price (P) times the quantity (Q) produced and sold. Total revenue increases as the quantity produced increases.
- Total costs increase with the quantity produced because it takes more inputs—such as workers—to produce more output.
- Marginal cost increases as more is produced because of diminishing returns to labor. As labor is increased with a fixed amount of capital, the increase in output from each additional unit of labor declines.

Profit Maximization and the Individual Firm's Supply Curve

profit maximization: an assumption that firms try to achieve the highest possible level of profits—total revenue minus total costs—given their production function.

To derive the firm's supply curve, we assume that the firm chooses a quantity of production that maximizes profits. This is the assumption of **profit maximization.** Now that we have seen how profits depend on the quantity produced, we can proceed to show how the firm chooses a quantity to maximize profits.

An Initial Approach to Derive the Supply Curve

Continuing with our pumpkin firm, we create a table that uses total revenue and total costs to calculate profits. Because total revenue depends on the price, we need a separate panel for each of three possible prices.

A Profit Table. Table 6.4 shows profits for your pumpkin firm. Total revenue, shown in the third column, increases with the number of pumpkins sold. Because total revenue depends on the price, we need a separate panel showing profits for each price. Table 6.4 has three panels, one for each of three possible prices. (The prices are the same as in Table 6.1.) Suppose the price of pumpkins is $35 a crate. Then if you sell 1 crate, the revenue is $35; if you sell 2 crates, the revenue is $70, and so on. Total revenue equals the price ($35) times the number of crates sold. Clearly, if no pumpkins are sold, the total revenue will be zero. Panels II and III show total revenue for two higher prices—$70 per crate and $100 per crate. For each total price, revenue increases as more is produced and sold.

Table 6.4 also shows how total costs increase with production. This is the same information already presented in Table 6.2. We repeat it here so that total costs can easily be compared with total revenue to calculate profits. Note how total costs, like total revenue, increase with production, from $50 for no pumpkins to $350 for 5 crates of pumpkins. The range of total costs is the same for all these panels because total costs do not depend on the price.

Table 6.4
Profit Tables Showing Total Costs and Total Revenue at Different Prices

PANEL I
If price equals $35 per crate, then production equals 2 crates.

Crates	*Total Costs*	*Total Revenue*	*Profits*
0	50	0	−50
1	70	35	−35
2	100	70	−30
3	150	105	−45
4	230	140	−90
5	350	175	−175

PANEL II
If price equals $70 per crate, then production equals 3 crates.

Crates	*Total Costs*	*Total Revenue*	*Profits*
0	50	0	−50
1	70	70	0
2	100	140	40
3	150	210	60
4	230	280	50
5	350	350	0

PANEL III
If price equals $100 per crate, then production equals 4 crates.

Crates	*Total Costs*	*Total Revenue*	*Profits*
0	50	0	−50
1	70	100	30
2	100	200	100
3	150	300	150
4	230	400	170
5	350	500	150

The last column of Table 6.4 shows profits: Total revenue minus total costs. Consider the $35 per crate price in panel I first. When no pumpkins are produced, profits are −$50. You lose $50 because $50 is paid for the land and total revenue is zero. If you produce 1 crate, the loss is $35; in other words, profits are −$35; the total cost of producing 1 crate ($70) minus the revenue from 1 crate ($35) equals −$35. For 2 crates, profits are still negative. Total revenue is $70 while total costs are $100, leaving a loss of $30. Three crates of pumpkins yield a loss of $45.

A glance down the last column in panel I shows that profits are negative at all production levels. In this case, any production at all may seem fruitless. But remember that you already paid $50 for the use of the pumpkin patch. Hence, it is best to produce 2 crates and cut your losses to $30. You still lose, but not as much as by producing only zero or 1 crate. This may seem strange because profits are negative.

But the profit-maximizing level of production is 2 crates. The maximum of profits would be −$30. Stated differently, the minimum loss would be $30.

The same type of profit-maximizing exercise with a different price is illustrated in panel II of Table 6.4. Here the price of pumpkins is $70, and so the total revenue is higher. If you sell nothing, then total revenue is zero and the loss is $50. If you sell 1 crate, total revenue is $70 and profits are zero. But if you sell 2 crates, total costs are $100 and total revenue is $140. Finally, some positive profit can be seen. But profits can be increased further: The profit-maximizing level of production is 3 crates.

Panel III shows profits for a higher price, $100 a crate. At this price, you would produce 4 crates. Profits would be $170. More or less production would lower profits.

In these three cases, you maximize profits by adjusting the quantity supplied. As the price rises from $35 to $70 to $100, the profit-maximizing quantity of pumpkins supplied goes from 2 crates to 3 crates to 4 crates. Thus, the price and the quantity supplied are positively related. This is the positively sloped supply curve.

■ **A Profit Graph.** The relationship between profits and production for your pumpkin firm given in Table 6.4 can be illustrated with a graph that compares total costs and total revenue. This is done in Figure 6.6. The curved line in the top graph of

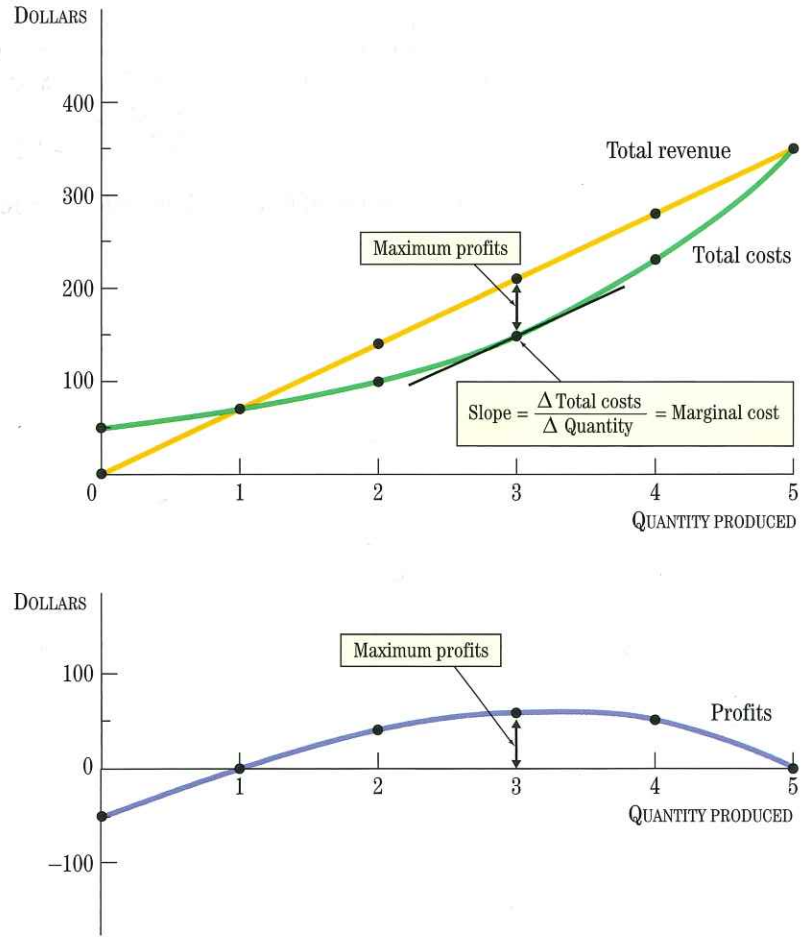


Figure 6.6
An Initial Approach to Profit Maximization
The top panel shows total costs and total revenue for a price of $70 per crate of pumpkins. Profits are the gap between total revenue and total costs. At the profit-maximizing point, the slope of the total costs curve equals the slope of the total revenue line. Hence, price equals marginal cost. The bottom panel shows explicitly how profits first increase and then decrease as production increases.

Figure 6.6 is the total costs curve. It corresponds to the total costs listed in Table 6.4 and is the same as the total costs curve in Figure 6.4. The upward-sloping straight line shows what total revenue would be for a price of $70 per crate. This line corresponds to the total revenue column in panel II of Table 6.4.

Profits are given by the gap between the total revenue line and the total costs curve. The gap—profits—is plotted in the lower panel of Figure 6.6. Note how profits first increase and then decrease as more is produced. The profit-maximizing firm chooses the quantity to produce that leads to the biggest gap, or the biggest level of profits. That quantity is 3 crates of pumpkins.

The Marginal Approach to Derive the Supply Curve

Economists use a different approach to analyzing profit maximization and deriving the supply curve. Once you know this approach, you will find it easier and faster than using total revenue, total costs, and total profits.

Continuing with our pumpkin firm, we first plot the marginal cost from Table 6.3 in Figure 6.7. Focus for now on the black dots in Figure 6.7; we derive the lines in the next few paragraphs. Each dot in Figure 6.7 represents the marginal cost of producing pumpkins at a different level of production. Figure 6.7 summarizes all that we need to know in order to find the quantity that a profit-maximizing firm will produce.

■ **Finding the Quantity Supplied at Different Prices.** Suppose the price of pumpkins is $10 a crate. Mark this point on the vertical axis of the diagram in Figure 6.7 with an arrow. We are going to derive the supply curve for your pumpkin firm by gradually raising the price from the low $10 value and determining how much you would produce at each price. At $10, the price is less than the marginal cost of producing 1 unit, which is $20, according to Figure 6.7. Would it make sense to produce pumpkins at this price? No, because producing 1 crate of pumpkins has a marginal cost of $20. The *additional* revenue that comes from producing one more

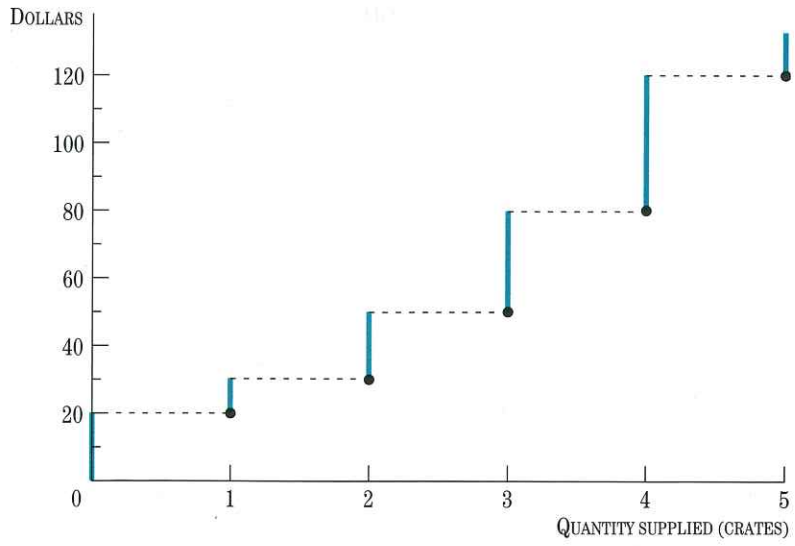


Figure 6.7
Derivation of the Individual Firm's Supply Curve
The dots represent the marginal cost from Table 6.3. At each dot, price equals marginal cost. These dots and the thick vertical lines indicate the quantity the firm is willing to supply at each price. Along the vertical lines, the firm produces the quantity that keeps marginal cost closest to price without exceeding it.

marginal revenue: the change in total revenue due to a one-unit increase in quantity sold.

crate is $10. The additional, or extra, revenue that results from producing and selling one more unit of output is called **marginal revenue.** Because laying out $20 and getting back $10 reduces profits, you would not bother to produce any pumpkins. In other words, the marginal cost of increasing production from 0 to 1 crate would be greater than the marginal revenue from selling 1 crate. Producing nothing would be the profit-maximizing thing to do.

Suppose the price of pumpkins rises. Move your arrow up the vertical axis of Figure 6.7. As long as the price is below $20, there is no production. Thus, the amount supplied at prices from $0 to $20 is given by the thick line at the bottom of the vertical axis, where quantity supplied equals zero.

Suppose the price rises to $20. Now the price equals the marginal cost, and the additional, or marginal, revenue from selling a crate of pumpkins will just cover the marginal cost of producing the crate. You now have sufficient incentive to produce some pumpkins. Strictly speaking, the price would have to be a little bit greater than $20 (say, $20.01) for you to earn more by producing 1 crate rather than 0 crates. At a price of exactly $20, you might be indifferent between 0 crates and 1 crate. At a price of $19.99, you would definitely produce nothing. At a price of $20.01, you would definitely produce 1 crate. The price of $20 is right between, but let's assume that you produce 1 crate rather than 0 crates at a price of $20. We indicate this in Figure 6.7 by showing that the quantity supplied is given by the black dot at 1 crate and $20.

Now consider further increases in the price. At prices above $20 up to $30, you would produce 1 crate because the price received for producing an extra crate is less than the marginal cost of $30. However, at a price of $30, the quantity supplied increases to 2 crates because price just equals the marginal cost of increasing production from 1 to 2 crates. A supply curve is now beginning to take shape. You can complete the curve by continuing to raise the price and watching what happens.

To shorten the story, let us move toward the other end of the scale. Suppose the price of pumpkins is $100. At $100, the price is greater than the marginal cost of producing the fourth crate, which is $80. Suppose you are producing 4 crates of pumpkins. Would it make sense to produce another crate? No, because increasing production from 4 crates to 5 crates has a marginal cost of $120. The marginal revenue that comes from producing one more crate is $100. Because laying out $120 and getting back $100 is a losing proposition, you would not do it. Production would stay at 4 crates. If production went up to 5 crates, profits would go down because the marginal cost of producing the fifth crate is greater than the marginal revenue. Producing 5 crates would not be a profit-maximizing thing to do.

What happens if the price rises to $110? At $110 the marginal cost is still less than the price, so it still makes sense to produce 4 crates. What if the price rises to $120? Then the price just equals the marginal cost at 5 crates, and you would produce 5 crates. When the price rises to the marginal cost at five units, then production increases to five units.

We have traced out the complete *individual* supply curve for your firm using Figure 6.7 with the assumption of profit maximization and the concept of marginal cost. The supply curve in Figure 6.7 is steplike; it consists of small vertical segments shooting up from the dots. Strictly speaking, it is only at the dots that price equals marginal cost. On the vertical segments above the dots, the price is actually greater than the marginal cost of production, but the price is not great enough to move on to a higher level of production.

In reality, however, for most products it is possible to divide production into smaller units—half crates, quarter crates, even a single pumpkin. As we do so, the jaggedness of the diagram disappears, as shown in Figure 6.8. It is the simple numerical example with production limited to whole-crate amounts that leads to steps in the supply curve. In reality, the diagram would consist of hundreds of dots rather

Green Pricing and Incentives

Profit maximization shows explicitly *why firms respond to incentives*. When the price of a good rises, a firm can increase its profits by producing more of that good. So the firm responds by producing more. The importance of incentives for firms has not gone unnoticed by practical people who have learned some economics. For example, the environmental group Friends of the Earth was looking for an "energetic person" to take a job as an "economics incentives associate" with responsibility for looking for "incentives for environmental protection."

Environmental protection is an area with fascinating and rapidly growing opportunities to use economic incentives. One of the new ideas is *green pricing*, used to encourage electric power firms to produce electricity using renewable resources, such as solar power or wind power. The goal is to reduce global warming and pollution caused when electricity is produced with coal or oil rather than with the sun or the wind.

Green-pricing programs have already been tried in many cities and states. The programs are very similar. Here is how the program works in Traverse City, Michigan. People agree to pay Traverse City Power and Light (the firm that produces electricity) an extra $7.50 per month (on average) if they can be sure that their electric power is produced with a wind turbine, rather than with a nonrenewable resource. This higher price is enough to cover the higher marginal cost of using wind power and thus gives Traverse City Power and Light the incentive to produce more electricity with wind power. Thus, a higher price for wind-produced electricity leads to a larger amount of wind-produced electricity. Price incentives work once again.

While green-pricing programs focus on firms' incentives (as in this chapter), the analysis in Chapter 5 suggests that such programs will have to focus more on consumers' incentives. People cannot actually tell the difference between wind-produced electricity and coal-produced electricity (both work just as well to power a VCR or a reading light), so they have no incentive to pay the extra amount. They have to volunteer. Perhaps concern for the environment will help get a sizable group of volunteers; Traverse City was considering producing stickers that people could put in their window to show that "they gave." But critics of green pricing argue that such efforts will never be large enough to have any noticeable effect on global warming or pollution. What do you predict? Will green-pricing programs continue to be small, or could they develop into a large program with a significant effect? Answering that question might be your first assignment if you accepted that job as "economics incentive associate."

than five dots. With hundreds of dots, the vertical segments would be too small to see and the firm's supply curve would be a smooth line. Price would equal marginal cost at every single point.

■ **The Price Equals Marginal Cost Rule.** In deriving the supply curve with Figure 6.7, we have discovered the key condition for profit maximization for a firm in

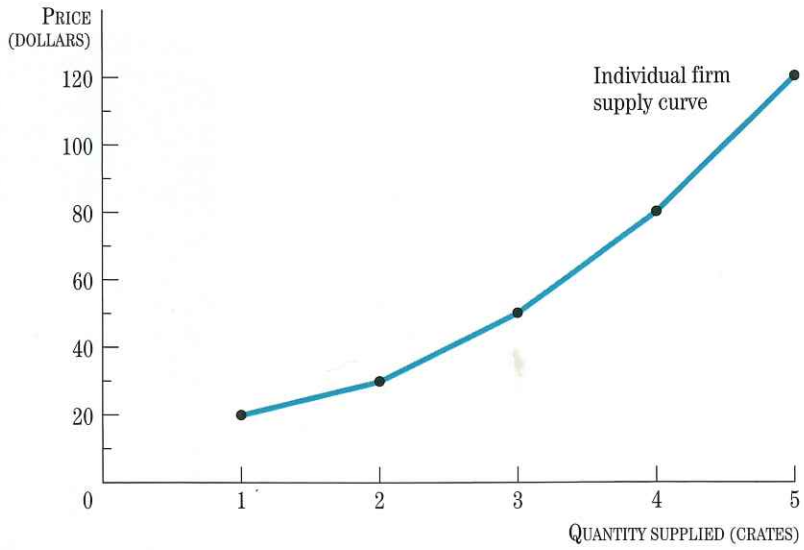


Figure 6.8
A Smooth Individual Supply Curve
If the firm can adjust its production by small amounts, the supply curve becomes a smooth line, as in this figure, rather than a series of steps, as in Figure 6.7. In some cases, such as the building of an airport, a dam, or a suspension bridge, fractions are not possible, and the supply curves will still have steps.

a competitive market: *The firm will choose its quantity such that price equals marginal cost.* You can see that from Figure 6.7. When the price is $80, the firm chooses a level of production for which the marginal cost equals $80 and produces 4 crates.

The price equals marginal cost rule for a competitive firm is a special case of a more general profit-maximization rule that we used, without calling it a rule, in our derivation of the supply curve. This more general rule is that *the firm will choose a quantity to produce so that marginal revenue equals marginal cost.* This more general rule makes intuitive sense for any profit-maximizing firm, whether it is a competitive firm or a monopoly. If the marginal revenue from producing an additional quantity of output is greater than the marginal cost, then the firm should produce that quantity; by doing so, it will increase total revenue by more than it increases total costs, and therefore it will increase profits. However, if the marginal revenue from an additional quantity is less than the marginal cost, then the firm should not produce that quantity. The firm maximizes profits by choosing the quantity of production for which marginal revenue equals marginal cost. Why is the price equals marginal cost rule a special case of the marginal revenue equals marginal cost rule? Because *for the case of a price-taking firm in a competitive market, the marginal revenue is equal to the price.* For example, as we showed above, if the price of pumpkins is $10 per crate, then the marginal, or additional, revenue from producing 1 crate of pumpkins is $10. Later, in Chapter 10, we will show that for a *monopoly,* the marginal revenue does not equal the price, so that even though marginal revenue equals marginal cost, the price does not equal marginal cost.

A Comparison of the Two Approaches to Profit Maximization

We have now considered two different approaches to profit maximization. One approach looks at the explicit relationship between profits and production. The other approach compares the price to the marginal cost. Both approaches give the

same answer. How do the approaches compare?

In Table 6.4 we looked at several prices, and we derived the profit-maximizing level of production by looking at profits for different levels of production at these prices. To do so, we had to create a new table for each price. This is quite time-consuming. In contrast, with the marginal cost approach, we only had to look at marginal cost for each unit of production and compare it with the price. Thus, the price equals marginal cost approach is considerably easier. Moreover, because marginal cost increases as the number of items produced increases, the price equals marginal cost approach tells us why the supply curve slopes upward. It is for these two reasons that economists usually use the price equals marginal cost approach.

REVIEW

- Profit-maximizing firms respond to higher prices by increasing the quantity they are willing to produce. Their response is the supply curve.
- The supply curve can be derived by comparing price and marginal cost. A profit-maximizing firm will produce a quantity that equates price and marginal cost.
- The upward-sloping marginal cost curve implies an upward-sloping supply curve.
- The supply curve can also be derived by looking at the relationship between profits and production. The profit-maximizing quantity is the same as that determined by the price equals marginal cost approach.

The Market Supply Curve

The *market* supply curve can be obtained by adding up the supply curves of all the *individual* firms in the market. Figure 6.9 gives an example in which there are two individual firm supply curves for pumpkins: One curve corresponds to your pumpkin firm, and the other, which is identical to yours, corresponds to the firm of your competitor, Fred, who is growing pumpkins on the other side of town. You and Fred have the same marginal cost for pumpkin growing, so your supply curves are exactly the same. You will both choose to produce the same number of pumpkins if the price is the same.

If only you and Fred are in the market, the market supply curve is the sum of just your two supplies. You get the market curve by adding in the horizontal direction, as shown in Figure 6.9. For example, if the price is $30, the quantity supplied by Fred will be 2 crates, and the quantity you supply will be 2 crates; thus the quantity supplied in the market at $30 is 4 crates. If the price rises to $50, Fred will produce 3 crates and you will also produce 3 crates; thus the quantity supplied in the market rises to 6 crates.

In reality, of course, there are more than two firms in a competitive market, and the individual supply curves for different firms in the market are usually different. But the concept of deriving the market supply curve is the same whether there are only 2 firms or 2,000 firms, and whether they are all the same or are all different. After adding up the individual supply curves for all the firms in the market, we arrive at a market supply curve like Figure 6.1. Thus, we have fulfilled one of the objectives of this chapter—deriving the market supply curve.

If there are many different firms in the market, the market supply curve can be much smoother than the individual supply curves. For example, a novelist may be

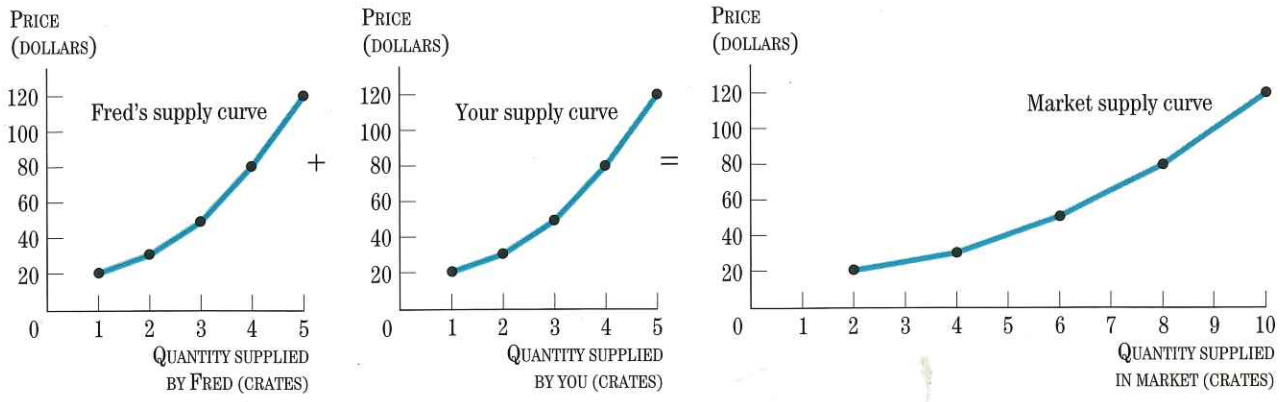


Figure 6.9
Derivation of the Market Supply Curve
The market supply curve is the sum of the individual firms' supply curves for all the firms in the market. The figure shows how the supply curves of two firms—Fred's and yours—sum to a market supply curve.

able to write only one novel a year. If a publisher offers a price to write the novel that is above the marginal cost of producing the novel, then the novelist will write the novel. Otherwise, the novel will not be written. But the market for novels in any one year consists of many authors with many different marginal costs. As the price of novels rises, more and more authors will decide to write novels, and the market supply curve for novels will look very smooth.

The Slope of the Supply Curve

We have shown that the slope and position of the individual firms' supply curves depend on the marginal cost at the different firms. If marginal cost rises very sharply with more production, then the supply curve will be very steep. If marginal cost increases more gradually, then the supply curve will be flatter.

Because the market supply curve is the sum of the individual firms' supply curves, its slope will also depend on marginal cost. The market supply curve can get very steep at high levels of production because marginal cost gets very high when production is high.

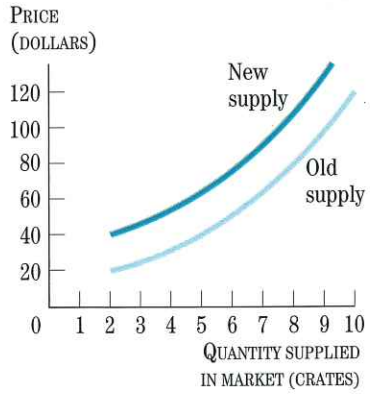


Figure 6.10
Shifts in the Market Supply Curve
An increase in marginal cost would shift the supply curve upward or to the left.

Shifts in the Supply Curve

Because the supply curve for the individual firm is given by its marginal cost, anything that decreases marginal cost will shift down the individual supply curves and therefore the market supply curve. For example, a new technology might reduce the marginal cost at every level of production. If this happens, then the market supply curve will shift down by the amount that marginal cost declines. Observe that a downward shift of a supply curve is equivalent to a rightward shift. Similarly, an increase in marginal cost—perhaps because of a disease affecting the pumpkins, so that more labor is required for each crate of pumpkins—would shift the supply curve upward or to the left (see Figure 6.10).

REVIEW

- The market supply curve is derived by adding up the individual supply curves of all the firms in the market.
- When the price rises, the individual firms in the market increase the quantity supplied. Hence, the market supply curve is upward-sloping.
- The slope of the supply curve depends on how sharply marginal cost increases.
- Anything that raises or lowers marginal cost will shift the market supply curve.

Producer Surplus

producer surplus: the difference between the price received by a firm for an additional item sold and the marginal cost of the item's production; for the market as a whole, it is the sum of all the individual firms' producer surpluses, or the area above the market supply curve and below the market price.

A firm would not produce and sell an item if it could not get a price at least as high as the marginal cost of producing that item. The **producer surplus** is the difference between the marginal cost of an item and the price received for it. For example, your marginal cost of washing cars on the weekend might be $4 per car. If the car-washing price in your area is $9 per car and this is the price you receive, then your producer surplus would be $5 per car. Or suppose, in your pumpkin firm, that the price of pumpkins is $25. Then you get $25 for producing 1 crate and incur $20 in marginal cost. The difference, $5, is your producer surplus. If the price of pumpkins is $35, you produce 2 crates and your producer surplus is $15 ($35 − $20) for the first crate plus $5 ($35 − $30) for the second crate, for a total of $20 producer surplus.

A Graphical Representation of Producer Surplus

The producer surplus can be represented graphically as the area above the individual firm supply curve and below the price line, as illustrated in Figure 6.11. The producer

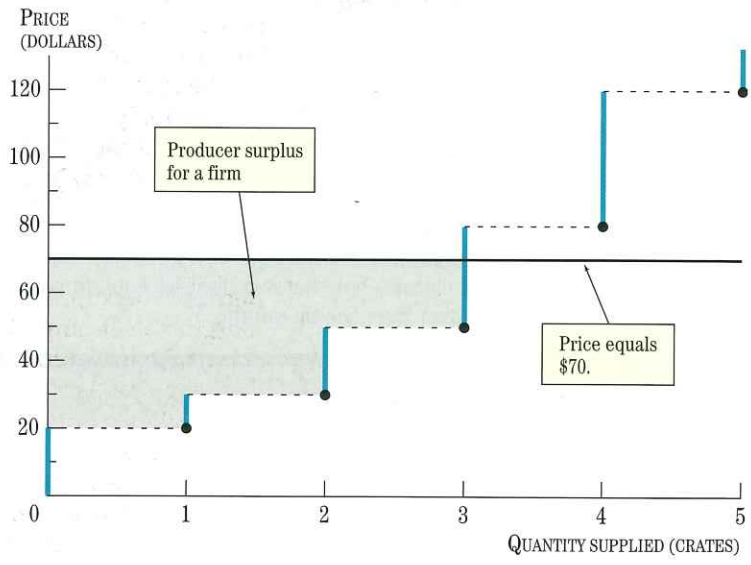


Figure 6.11
Producer Surplus for an Individual Firm
As shown here, for an individual firm, the producer surplus is the area between the price line and the supply curve.

READING THE NEWS ABOUT... Supply Shifts Caused by Natural Disaster

Marginal cost is defined as the increase in total costs associated with each additional unit of production. Anything that decreases marginal cost will shift the supply curve down (or to the right), and anything that increases marginal cost will shift the supply curve up (or to the left). Reading the article below about how U.S. firms have adjusted to the effects of the tsunami, can you identify the firms that are likely to see a shift in their supply curve in response to increased marginal costs—and those that are likely not to experience such a shift?

After Tsunami, US Firms Adjust

by Diane E. Lewis
Global Staff
January 6, 2005

US companies whose goods are made in Southeast Asia are shifting production and changing distribution routes in order to limit the commercial impact of the deadly earthquake and tsunami.

Clothing retailer Gap Inc., with more than 700 manufacturing contractors worldwide, sought to minimize disruptions by transferring shipments bound for an Indian port to another site in the country.

"There were problems in southern India," said spokeswoman Kris Marubio. "Roads leading to the major thoroughfare were impacted. So we moved them to other ports in India. We did this over the past few days."

Although Southeast Asia has become a hub for apparel and footwear manufacturers, big US companies have so many contractors that they can readily shift production from one country to another without affecting retail prices, analysts say.

Gourmet coffee roasters, however, don't have that luxury. The natural disaster could limit the supply of coffee grown on the island of Sumatra and drive up already high prices.

Chuck Coffman, president of Armeno Coffee in Northborough, said Sumatran coffee has risen to about $3 per pound, almost double what he had been paying last year. He sells 17 varieties of Indonesian coffee, including Sumatran coffee made from the Arabica bean, prized among gourmands.

But this week, Coffman learned that future supplies could be jeopardized by lack of workers and bad roads.

"No one knows what will happen," said Coffman. "What we are fearing is that the crop will mature, but there will not be enough workers to pick it, and they may not be able to get it down from the mountains."

surplus is analogous to the consumer surplus, the area below the demand curve and above the price line, derived in Chapter 5.

The producer surplus in the whole market can be obtained by adding up the producer surplus for all producers or by looking at the area above the market supply curve and below the price. This is illustrated in Figure 6.12.

Before the natural disaster, the price of Sumatran coffee had risen 25 cents a pound last year because of concerns over limited supply. Now consumers will be paying more for Sumatran coffee.

"Potentially, I think there will be a higher acceptance for the higher price by the consumer," said Thomas Fricke, chief executive and cofounder of organic coffee producer ForesTrade, which gets 60 to 70 percent of its coffee from the Aceh region, one of the hardest-hit areas.

Consumers "have the perception that people in Indonesia are having a difficult time and may be willing to accommodate higher prices, regardless whether the impact warrants it or not," he said.

Consumers could also see tuna and other fish prices increase. The tsunami destroyed fishing fleets in Indonesia, India, and Thailand, said John Connelly, president of the National Fisheries Institute in Washington, D.C. "Fishing fleets have been decimated, and that could impact the tuna we get," he said, adding that Thailand is a major exporter of tuna to the United States.

UPS, the global package delivery service, dealt with the disaster by implementing a plan to get relief and commercial goods into parts of Indonesia with smaller aircraft.

"Commerce is pretty much at a standstill in some areas because there is so much congestion due to the aid relief in the airport and ports," said spokesman John Flick. "Before the tsunami, when we flew to Jakarta, we used one aircraft. Now, Jakarta has been taken out of the equation. We're going directly to the smaller islands, and going directly to smaller aircraft to move goods."

While many apparel and footwear factories may have escaped damage, distribution problems could arise, analysts say. Footwear companies like Reebok and Nike, which produce a good portion of their shoes in the region, could face some obstacles if they rely on trucks to transport goods along roads uprooted by the earthquake, or use ports glutted with goods from relief efforts.

"The biggest impact may be that some of the trucks that move product from the factory to ports will not be available," said Madison Riley, a principal and national service director for consulting firm Kurt Salmon & Associates Inc.

"Future distribution could be impacted somewhat," he said. "But I don't see prices going up. Instead, if products are delayed retailers might cancel some of it and wholesalers could be stuck with it. But there could be a greater opportunity for lower-cost products."

At Reebok, where about 40 percent of its shoes are made in Thailand and Indonesia, there were no disruptions to the firm's ***supply chain*** or production, said spokesman Denise Kaigler.

Nike said one of its apparel contractors in Sri Lanka appears to have been damaged by the tsunami, but that would not affect overall production. Five percent of the firm's global apparel production is located in the region, said Caitlin Morrs, Nike's senior manager of global issues management.

As of October, Nike had 42 contractors in Indonesia, including nine footwear producers, 29 apparel makers, and five producers of athletic equipment such as yoga mats and golf bags, she said. Most were based in Jakarta, where there is no damage.

The applications of producer surplus are similar to those of consumer surplus. Producer surplus provides a measure of how much a producer gains from the market. The sum of producer surplus plus consumer surplus is a comprehensive measure of how well a market economy works, as we will see in Chapter 7.

What Is the Difference Between Profits and Producer Surplus?

Profits and producer surplus are not the same thing. Profits are the difference between total revenue and total costs, while the producer surplus measures the difference between the price and the marginal cost of every unit. How can we compare these two measures?

Suppose the price of pumpkins is \$70 per crate; then you are willing to produce 3 crates of pumpkins. Total revenue is \$210 and total costs are \$150; thus, you are making a \$60 profit. (See panel II of Table 6.4.) How much is your producer surplus when 3 crates are sold at \$70? As just defined,

$$\begin{aligned}\text{Producer surplus} &= (P - MC_1) + (P - MC_2) + (P - MC_3)\\ &= (\$70 - \$20) + (\$70 - \$30) + (\$70 - \$50) = \$110\end{aligned}$$

where MC_1 is the marginal cost of the first crate, MC_2 is the marginal cost of the second crate, and MC_3 is the marginal cost of the third crate. Thus, profits are \$60 and producer surplus is \$110.

Notice that there is a difference of \$50 between profits and producer surplus. That number happens to be equal to the fixed costs of the firm. Thus, producer surplus equals profit plus fixed costs. Try the same method for different prices and quantities sold, and you will arrive at the same result: Producer surplus always equals profits plus fixed costs. We now show that this is no coincidence.

Consider Table 6.5. The first four columns of Table 6.5 are already familiar from the earlier tables. The fifth column shows the sum of the marginal costs for all that is produced. For example, the sum of marginal costs when production is 3 crates is \$20 plus \$30 plus \$50 equals \$100. The last column shows the difference between total costs and fixed costs—that is, variable costs. Notice that the last two columns are equal.

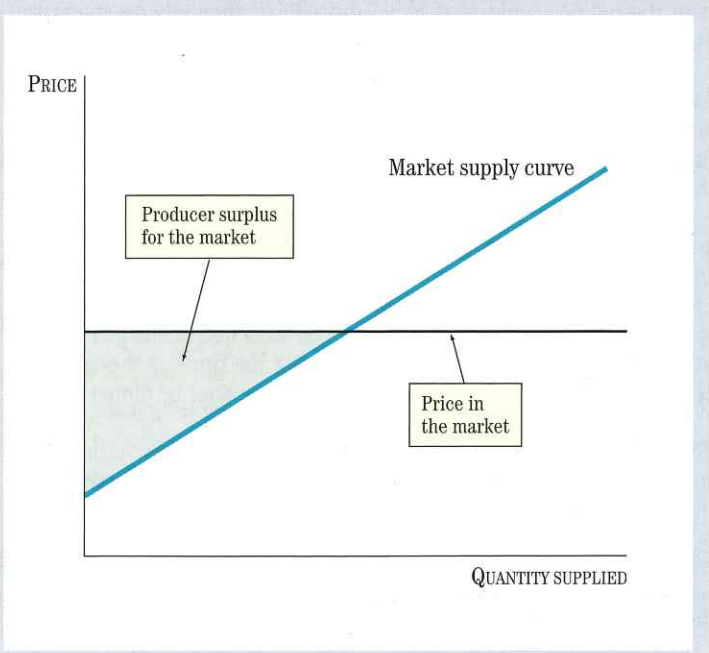


Figure 6.12
Producer Surplus for the Market
If we add up producer surplus for every firm, we get the producer surplus for the whole market. This is given by the area between the price and the market supply curve.

Thus, as we sum up marginal costs for any quantity Q produced, we count all costs except fixed costs. In other words,

Sum of marginal costs = total costs − fixed costs = variable costs

When the producer sells a quantity Q, we can say that

$$\text{Producer surplus} = (P - MC_1) + (P - MC_2) + (P - MC_3) + \cdots + (P - MC_Q)$$

The number of terms in this sum is Q. For the example above, $Q = 3$ and there were Q terms in the sum. Thus, we can translate this definition of producer surplus into the price (P) of the good times the quantity (Q) sold minus the sum of the marginal costs of all units. That is,

$$\text{Producer surplus} = (P \times Q) - \text{sum of marginal costs}$$

As we now know, $P \times Q$ is the total revenue, and the sum of marginal costs equals the difference between total costs and fixed costs. Substituting these relationships into the preceding gives

Table 6.5
Summing Up Marginal Costs

Crates of Pumpkins	Fixed Costs	Total Costs	Marginal Cost	Sum of Marginal Costs	Total Costs Less Fixed Costs
1	50	70	20	20	20
2	50	100	30	50	50
3	50	150	50	100	100
4	50	230	80	180	180
5	50	350	120	300	300

Producer surplus = total revenue − (total costs − fixed costs)

Finally, since profits equal total revenue minus total costs, we obtain the relationship between producer surplus and profits:

Producer surplus = profits + fixed costs

REVIEW

- Producer surplus is the price a firm receives for selling a unit of a product minus the marginal cost of producing that unit.
- The producer surplus for a firm is the area below the price line and above the firm's supply curve.
- For all the firms in a market, the producer surplus is the area below the price line and above the market supply curve.
- Producer surplus is different from profits. Producer surplus is greater than profits by the amount of fixed costs.

Conclusion

In this chapter, we have derived the supply curve in a competitive market by looking at the behavior of firms. We assumed that a firm decides how much to produce by maximizing profits. The firm makes this decision taking prices as given and considering its production function, which relates the number of hours of work at the firm to the output of the firm. The production function enters the firm's profit calculations through its effects on the firm's costs. Because the production function has diminishing returns to labor, the firm faces increasing marginal cost. From the firm's marginal cost, we can quickly find the firm's supply curve. Profit maximization implies that the firm will produce the quantity where price equals marginal cost.

The connection between marginal cost and the supply curve is fundamental to understanding how markets work. We will make use of this connection many times throughout this book, especially when we consider public policy issues, such as the efficiency of markets, taxation, and regulation of firms. When economists see or draw a supply curve, they are usually thinking about the marginal cost of the firms that

underlie the supply curve. The supply curve and the marginal cost curve are virtually synonymous for economists.

The price equals marginal cost rule for a profit-maximizing firm is fundamental for understanding how well markets work. When, in Chapter 7, we combine this rule with the analogous rule that the price equals the marginal benefit of a good for a consumer, we will discover an attractive feature of competitive markets.

We have examined firm behavior in this chapter and consumer behavior in Chapter 5, and the next step in our analysis of markets is to examine the interaction of these firms and consumers. That is the subject of Chapter 7.

KEY POINTS

1. A firm is an organization that uses inputs to produce goods or services.
2. In competitive markets, firms are price-takers.
3. The foundations of supply are found in the profit-maximizing behavior of firms.
4. Profits are defined as total revenue minus total costs.
5. The production function shows how production increases with more labor; the marginal product of labor declines as more labor is added and capital is not changed.
6. Marginal cost increases as more is produced because of diminishing returns to labor.
7. A price-taking firm produces up to the point where price equals marginal cost, which is the key rule for profit maximization.
8. The reason the supply curve slopes upward is that marginal cost is increasing. A higher price enables the firm to produce at higher levels of marginal cost.
9. We can also determine the profit-maximizing quantity of production by looking at how profits depend on production and finding the highest level of profits.
10. The market supply curve is obtained by adding up the individual supply curves. The market supply curve can be smooth even if the individual supply curves are not.
11. Producer surplus is the area above either the individual or the market supply curve and below the price line.
12. Producer surplus and profit are not the same thing. Producer surplus equals profits plus fixed costs.

KEY TERMS

firm
price-taker
competitive market
profits
total revenue
total costs
production function
marginal product of labor
diminishing returns to labor
fixed costs
variable costs
marginal cost
profit maximization
marginal revenue
producer surplus

QUESTIONS FOR REVIEW

1. What is a firm?
2. What is the difference between a corporation and a sole proprietorship?
3. Why do total costs increase as more is produced?
4. Why does marginal cost increase as more is produced?
5. What is the relationship between an individual supply curve and marginal cost?
6. Why does profit maximization imply that price equals marginal cost?
7. Why would a firm never choose to produce at a point where the price of an item is less than the marginal cost of the item?
8. What does it mean to say that firms are price-takers?
9. When does it make sense to assume that firms are price-takers?
10. How is the market supply curve derived from individual supply curves?
11. Why might the market supply curve be smoother than the individual supply curves?
12. What is producer surplus?

PROBLEMS

1. The table at the top of the next page shows the total costs of producing strawberries on a small plot of land.

Pounds of Strawberries	Total Costs (dollars)
0	10
1	11
2	14
3	18
4	25
5	34

a. Calculate the marginal cost schedule.
b. Draw the farmer's supply curve.
c. Suppose the price of 1 pound of strawberries is $4. How much would this farmer produce? Show graphically the area of producer surplus. What are profits?
d. Suppose the price of strawberries goes up to $7 per pound. How much will the farmer produce now? What are profits now?

2. Consider the example of the cost of pumpkins in Table 6.4. Compute the total revenue, total costs, and profits when the price of a crate of pumpkins is $60. How many crates of pumpkins will maximize profits? Now find the profit-maximizing quantity by using the marginal cost approach. How do your answers compare? Which approach did you find easier?

3. Consider the following information:

Daily Production and Costs at Jill's Bread Bakers

Quantity Produced (dozens of loaves)	*Total Costs (dollars)*
0	20
1	22
2	26
3	32
4	40
5	50
6	62
7	76

a. Calculate the marginal cost for Jill's bread production.
b. Draw the supply curve for this firm.
c. Jill can sell as many loaves as she wants in the market at a price of $12 for a dozen loaves. How many loaves will she sell each day? Use your diagram to show how much producer surplus she receives.

4. Suppose you are able to mow lawns at $12 per hour. The only cost to you is the opportunity cost of your time. For the first 3 hours, the opportunity cost of your time is $9 per hour. But after 3 hours, the opportunity cost of your time rises to $15 per hour because of other commitments. Draw the marginal cost to you of mowing lawns. Draw in the price you receive for mowing lawns. For how long will you mow lawns? Calculate your producer surplus.

5. Suppose a price-taking firm has the following total costs schedule:

Quantity	Total Costs
0	20
1	30
2	42
3	55
4	75
5	100
6	130

a. Calculate marginal cost. If the price in the market is $20, how many units will the firm produce?
b. Suppose the price in the market falls to $12 per unit. How many units of output will this firm produce in order to maximize profits?
c. Suppose there is an improvement in technology that shifts total costs down by $8 at every level of production. How much will the firm produce and what will profits be at a price of $20 and at a price of $12?

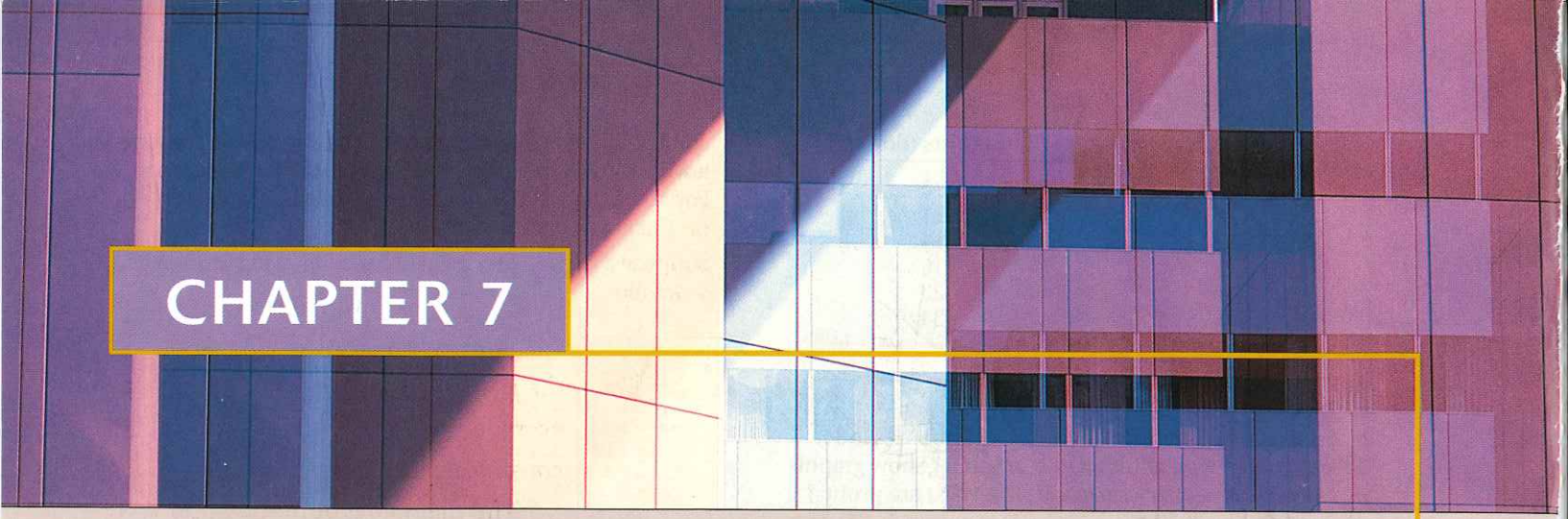
6. Consider the following information about a firm:

Quantity	Total Costs	Total Revenue
0	500	0
1	700	500
2	1,100	1,000
3	1,500	1,500
4	2,300	2,000
5	3,500	2,500

On the same diagram, plot the total revenues and total costs curves for this firm. What are the maximum profits that this firm can earn? Show this level of profits in the diagram. Do the slopes of the two curves appear to be the same at the maximum profit level?

7. Using the information in problem 6, find the fixed costs and the producer surplus when the firm produces the profit-maximizing quantity. Show that this amount of producer surplus equals profits plus fixed costs.

8. Does the assumption that firms are price-takers seem less valid than the assumption that consumers are price-takers? Explain why. Suppose that there are 50 firms and 50 consumers in the market. Would both assumptions be equally accurate in that case?



CHAPTER 7

The Interaction of People in Markets

This is an old but true story, going back to before the field of economics even existed. It is about an absent-minded philosophy professor who was interested in human interaction. He was particularly fascinated by how the economy, consisting of the interaction of millions of people pursuing their own interests, worked. He did not have much to go on; there were no economics professors at his school or at any other school. So, although he was a gifted teacher, he quit his teaching job at the university and traveled; he interviewed businesspeople; he visited factories; he talked to workers; he watched ships come and go; he studied the economies of other countries and of other times. He did everything he could to find about the economy. Amazingly, not only was he able to pull all this material together into a coherent view, he also managed to get it down on paper for other people to read, learn, and enjoy. By doing so, he invented the field of economics. His view of the economy is still dominant today.

The professor's name was Adam Smith, and the book he wrote, called *Wealth of Nations*, was first published in 1776. His deepest insight, among many deep insights, is called the **invisible hand** theorem, still most often stated using his words: "It is not from the benevolence of the butcher, the brewer, or the baker that we expect our dinner, but from their regard to their own interest." And whether it is the butcher, the brewer, or the baker, he is "led by an invisible hand to promote an end which was no part of his intention. . . . By pursuing his own interest he frequently promotes that of the society more effectually than when

Corporate leaders gather in a field outside Darien, Connecticut, where one of them claims to have seen the invisible hand of the marketplace.

he really intends to promote it." In other words, without any formal coordination, firms (butchers, brewers, bakers, and many others) that are pursuing their own interests interact with consumers who are also pursuing their own interests, and somehow everyone ends up producing and consuming a quantity that is efficient.

The main goal of this chapter is to state clearly and prove the invisible hand theorem. The theorem is not always true, and we want to be clear about the circumstances in which it is true. We first need to explain what is meant by *efficient* (the modern term for Smith's "effectual"), and then show why and under what circumstances the quantity produced and consumed is efficient. We also show how to measure the economic loss from producing more or less than the efficient quantity.

invisible hand: the idea that the free interaction of people in a market economy leads to a desirable social outcome; the term was coined by Adam Smith.

Chapters 5 and 6 have paved the way for our goal in this chapter. In Chapter 5 we studied consumers. We can say that consumers are pursuing their own interests, because they maximize their utility. In Chapter 6 we studied firms. We can say that firms are pursuing their own interests, because they maximize profits. Now we study the interactions of firms and consumers in competitive markets. Figure 7.1 is a schematic illustration of the model we use to explain this interaction and thereby explain the invisible hand theorem. The model, called the **competitive equilibrium model,** is an embellishment of the supply and demand model discussed in Chapter 3, but now with the behavior of consumers and firms explicit.

competitive equilibrium model: a model that assumes utility maximization on the part of consumers and profit maximization on the part of firms, along with competitive markets and freely determined prices.

Individual Consumers and Firms in a Market

In our analysis of economic interaction, it is very important to think about what individual consumers and firms are doing. Consider an example of consumers and producers of the same commodity: long-stemmed roses. Maria and Ken are two of many potential rose consumers who are deciding how many roses to buy. Both are willing to pay a certain amount for roses, but not necessarily the same amount. Hugo and Mimi are two of many rose producers who are deciding how many roses to produce in their gardens. Both have marginal costs for producing roses, but not necessarily the same marginal costs.

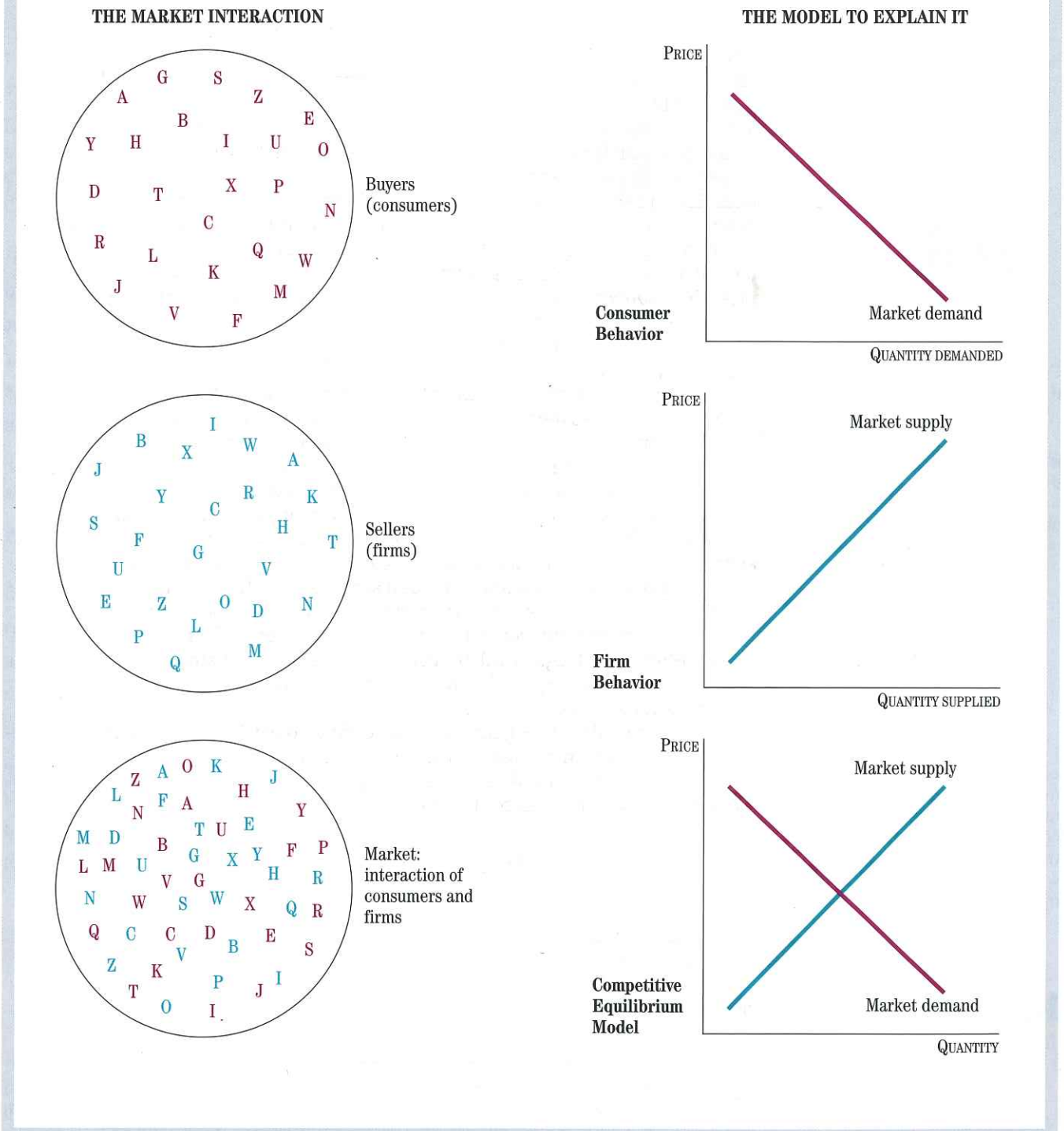


Figure 7.1
The Market Interaction and the Model to Explain It

In this chapter, we explain how individual buyers and sellers interact in a market (on the left) by combining the behavior of consumers (Chapter 5) with the behavior of firms (Chapter 6) to get a model of competitive equilibrium (Chapter 7).

The Hard Way to Process Information, Coordinate, and Motivate

The rose decisions of Maria, Ken, Hugo, Mimi, and all the others in the market clearly interact with one another. For example, an increase in Hugo's marginal costs—perhaps because of an extra rose-processing expense to ward off a new insect—will probably reduce the amount of roses he decides to produce; either this means less rose consumption for Ken, Maria, and other consumers, or it means more rose production for Mimi and other producers. Similarly, if Ken decides to purchase more roses, someone else must decide to decrease consumption or increase production. How are all these decisions worked out? What *information* is needed in order to determine whether it is better for Mimi's garden to produce more or for Hugo's garden to produce more? What *coordinates* a change in consumption or production by one person with an offsetting change in consumption or production by other people? What *motivates* some people to consume less and others to produce more if one person decides to consume more?

Suppose you had to work this out. To make your job easier, suppose that Maria, Ken, Hugo, and Mimi were the whole world as far as roses go. If you and they were all in one place together, you might imagine conducting their consumption and production activities the way the leader of a marching band would conduct the band members. You raise your baton toward Maria to signal more consumption; you shake your head at Ken to signal less consumption; you point your finger at Hugo to signal more production; you turn your back on Mimi to signal no change in production; you blow your whistle to signal when to begin consuming and producing.

To provide motivation, you might change your facial expression when you look at Maria; a frown or perhaps a smile may help to motivate her to purchase more roses. Your choice of which finger to point at Hugo may affect his motivation, and the shrill of your whistle might serve to motivate them all to do what you say.

To do your job right, you will also need to have information about rose production for Mimi's and Hugo's gardens. For example, to know whether it is appropriate to point your finger at Hugo and turn your back on Mimi, you need to know which garden has lower marginal costs of rose production.

If this is not already beginning to sound ridiculously impossible, remember that if you had this job in the real world, you would have to coordinate, motivate, and know intimately millions of consumers and producers. This is an amazingly complex job even for this single, relatively simple commodity.

The Easy Way to Process Information, Coordinate, and Motivate

Fortunately, you do not need to worry about being called upon to perform such an impossible task. There is a remarkable device that does the information processing, coordinating, and motivating for us. No one person invented this device; it evolved slowly over thousands of years and is probably still evolving. It is called *the market* (in this case, the long-stemmed-rose market). Of course, like many markets, the rose market does not take place in any one location. It consists of all the florists, street carts, and farmers' markets where roses are sold and all the gardens and greenhouses where roses are grown, whether in the United States, Europe, Latin America, Africa, Australia, or Asia. Fortunately, a market can serve as an information-processing, coordinating, and motivating device even if it does not take place at any one location. Buyers and sellers never have to see one another.

How does the market work? What will be the total quantity of roses consumed? Who will consume what amount? What will be the total quantity of roses produced?

One-Stop Shopping for Processing Information, Coordinating, and Motivating: The Market
This New York City flower vendor represents just one piece of the huge and multifaceted market for flowers that exists throughout the world.

Which garden will produce what amount? Let us see how economists answer these questions about how people interact in a market.

The Competitive Equilibrium Model

Economists use the *individual* demand curves and the *individual* supply curves derived in Chapters 5 and 6 to describe what happens to consumers and firms when they interact in a market.

Recall that each of the individual demand curves depends on the marginal benefit—the willingness to pay for additional consumption—the individual gets from consuming the goods. Together these marginal benefits create a market demand curve for roses. The demand curve shows how much consumers in total are willing to buy at each price.

Recall also that individual supply curves depend on the marginal costs of the firms. Together their marginal costs create a market supply curve for roses. The supply curve shows the total quantity supplied by all firms at each price.

The resulting market demand and supply curves are shown in the center of Figure 7.2, flanked by Maria's and Ken's individual demand curves and by Hugo's and Mimi's individual supply curves. Note that we have used the same units for the quantity supplied and the quantity demanded in Figure 7.2; the price (P, measured in dollars per rose) is on the vertical axis, and the quantity (Q, the number of roses) is on the horizontal axis. We continue to assume that Maria, Ken, Hugo, and Mimi are the whole market so that we can show the market in one diagram. A competitive market would typically require more buyers and sellers.

We have seen supply and demand curves like those in the center of Figure 7.2 before in Chapter 3. But now—after Chapters 5 and 6—we know much more about what the demand and supply curves mean. Individual consumer behavior and individual firm behavior are now seen as underlying the supply and demand model. To emphasize that the supply and demand model incorporates utility-maximizing consumers and profit-maximizing firms in competitive markets, we refer to it as the

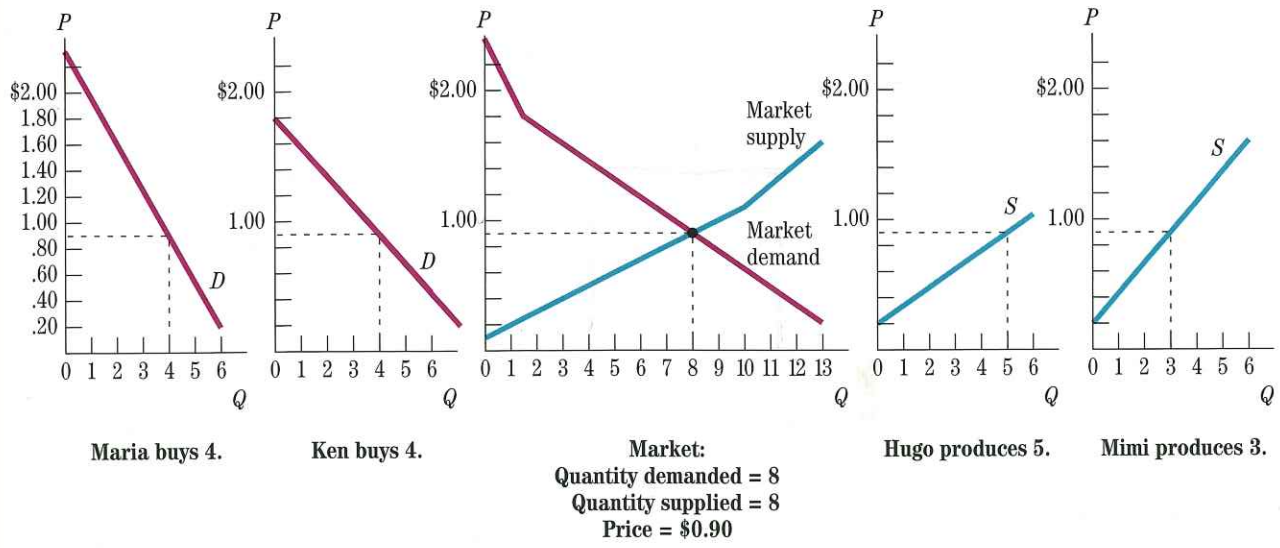


Figure 7.2
Price and Quantity Determination

The market demand curve is the sum of the individual demand curves of the consumers in the market. The market supply curve is the sum of the individual supply curves for the producers. When quantity demanded equals quantity supplied, the two curves intersect. The equilibrium price and quantity will be given by this intersection. The individual firm and consumer decisions can then be read off the individual supply and demand curves at those prices.

competitive equilibrium model. The competitive equilibrium model, as we have said, is simply the supply and demand model with the behavior of consumers and firms made explicit. Because the competitive equilibrium model has more to it than the supply and demand model, we can do more with it.

Individual Production and Consumption Decisions. A key prediction of the competitive equilibrium model is that a price will emerge from the interaction of people in the market such that the quantity supplied *equals* the quantity demanded. This is the *equilibrium price.* Graphically, the price is given at the point of intersection of the market supply curve and the market demand curve; here the quantity supplied in the market equals the quantity demanded in the market. For the example shown in Figure 7.2, the equilibrium price is $.90 a rose.

equilibrium price: the price at which quantity supplied equals quantity demanded. (Ch. 3)

Once we have determined the price in this way, the supply and demand curves tell us how much in total will be consumed and produced at that price. We look at the market demand curve and see how much is demanded at that price, and we look at the market supply curve and see how much is supplied at that price. Because the curves intersect at the market price, the quantity demanded and the quantity supplied are the same. They are at the point on the horizontal axis directly below the intersection. In Figure 7.2, the quantity bought and sold is 8 roses.

Thus far, we have not done anything more with the competitive equilibrium model than we did with the supply and demand model. But now, armed with the price, we can go to the individual demand curves to see how much Maria and Ken will buy. Look to the left in Figure 7.2 to find the quantity demanded by Maria and by Ken when the price is $.90 a rose. They each buy 4 roses. Maria and Ken are motivated to buy this amount—without any central coordinator—because, at $.90 a rose,

they maximize their respective utilities by consuming this amount. Observe that Maria and Ken do not have the same individual demand curves. Nevertheless, the quantity demanded by each can still be determined from their demand curves, as shown in Figure 7.2.

The individual supply curves tell us how much Hugo and Mimi will produce. Look to the right in Figure 7.2 to see how much Hugo and Mimi produce when the price is $.90 a rose. Hugo produces 5 roses, and Mimi produces 3 roses. Hugo and Mimi are motivated to produce this amount—again without any central coordinator—because, at $.90 a rose, they maximize their profits by producing this amount.

In sum, the competitive equilibrium model, which includes the behavior of the consumers and the firms, predicts the price, the quantity consumed by each person, and the quantity produced by each firm. It also predicts a certain marginal benefit of consumption for each consumer and a certain marginal cost for each producer. Hence, the model provides answers to all the questions posed earlier.

Adjustment to the Equilibrium Price. As can be seen from Figure 7.2, if the price is higher than the predicted market price at the intersection of the supply curve and the demand curve, then the quantity supplied is greater than the quantity demanded; we say that there is a *surplus*. When there is a surplus, the price will fall, resulting in demand increasing and supply decreasing until the surplus disappears. However, if the price is lower than the predicted market price, then the quantity demanded is greater than the quantity supplied; we say that there is a *shortage*. When there is a shortage, the price will rise, resulting in demand decreasing and supply increasing until the shortage disappears. At this low price, the marginal benefit is greater than the marginal cost, and the price will rise until the shortage disappears. Thus, if the price falls when there is a surplus and rises when there is a shortage, the price will converge to the equilibrium price.

surplus (excess supply): the situation in which quantity supplied is greater than quantity demanded. (Ch. 3)

shortage (excess demand): the situation in which quantity demanded is greater than quantity supplied. (Ch. 3)

REVIEW

- Centrally coordinating and motivating the thousands of consumers and producers of any good would be an amazingly complex task requiring a vast amount of information.
- The market is a device that provides information and coordinates and motivates consumers and producers in a decentralized way. The market does this job in a way that no one individual can.
- Economists describe the interactions of people in the market through the competitive equilibrium model. According to the model, the equilibrium price and total quantity are given by the intersection of the market supply and demand curves; individual decisions about consumption and production are given through the individual demand and supply curves, which are based on utility maximization and profit maximization.

A Double-Auction Market

How well does the competitive equilibrium model work in explaining the actual interaction of individual firms and consumers? Economists answer this question by observing markets in which one can see exactly what all the buyers and sellers do. Because it is difficult, if not impossible, to observe all the participants in actual markets, it is necessary to set up experimental markets for this purpose. An experimental

market is much like a real-world market except that one can observe all the actions of the participants. In this section we describe such a market and compare it to the model.

double-auction market: a market in which several buyers and several sellers state prices at which they are willing to buy or sell a good.

Market Participants and Their Incentives

A simple kind of market in which several buyers and several sellers interact is the **double-auction market.** In a double-auction market, *both* buyers and sellers call out prices. Buyers bid a certain price for items they want to buy, and sellers ask a certain price for items they want to sell. Four or five buyers and four or five sellers are enough to make the market work, but many more can also participate.

Many real-world markets are like double-auction markets. For example, the New York Stock Exchange and the commodity exchanges in Chicago are double-auction markets. In Chicago, traders in trading pits call out whether they want to buy and sell.

As soon as the market opens, buyers can bid and sellers can ask certain prices. A transaction takes place any time a buyer accepts a price a seller asks or a seller accepts a price a buyer bids.

In order to make this experimental market work like markets in the real world, the buyers and sellers must be given some objectives and take their actions seriously. Each buyer is given a marginal benefit schedule. Each seller is given a marginal cost schedule. A small sheet of paper describing these marginal benefits or marginal costs is given to each of the buyers and sellers who participate in the market. An example of both a seller's marginal cost sheet and a buyer's marginal benefit sheet is shown in Table 7.1.

Table 7.1
Marginal Benefit and Marginal Cost for a Double-Auction Market

Example Buyer Sheet

Number of Items	*Marginal Benefit (dollars)*
1	25
2	20
3	15
4	10
5	5

Example Seller Sheet

Number of Items	*Marginal Cost (dollars)*
1	1
2	6
3	11
4	16
5	21

■ **Buyers Earn a Consumer Surplus.** During each trading period, buyers may purchase any number of items but can bid for only one unit at a time. For each item successfully purchased, the buyer receives the amount listed on the sheet under the column marked "marginal benefit." Thus, the buyer's personal gain on each item purchased is the difference between the marginal benefit of that item and the amount paid for it. Notice that a motivated buyer will want to get the lowest price possible. Other buyers will be competing to do the same thing. It would not be wise for a buyer to buy an item for more than the marginal benefit because that would result in a loss.

consumer surplus: the difference between what a person is willing to pay for an additional unit of a good—the marginal benefit—and the market price of the good. (Ch. 5)

For example, suppose you have the buyer's sheet shown in Table 7.1 and you buy two items. The marginal benefit from the first item is \$25, and the marginal benefit from the second item is \$20. If you pay \$15 for the first item and \$10 for the second item, then your total gain is (\$25 − \$15) + (\$20 − \$10) = \$20. Observe that each of the terms in this sum is the marginal benefit minus the price, or the *consumer surplus.*

Be sure to distinguish between *market surplus* and *consumer surplus* or *producer surplus.*

■ **Sellers Earn a Producer Surplus.** During each trading period, sellers are free to sell any number of items but can ask a price for only one item at a time. Each item sold costs the amount listed on the sheet under the column marked "marginal cost." The seller's personal gain on each item sold is the difference between the price the item sold for and the marginal cost of the item. Notice that a seller is motivated to get the highest price but is competing with other sellers who may be asking lower prices. It would not be wise for a seller to accept a price lower than the marginal cost, because that would result in a loss.

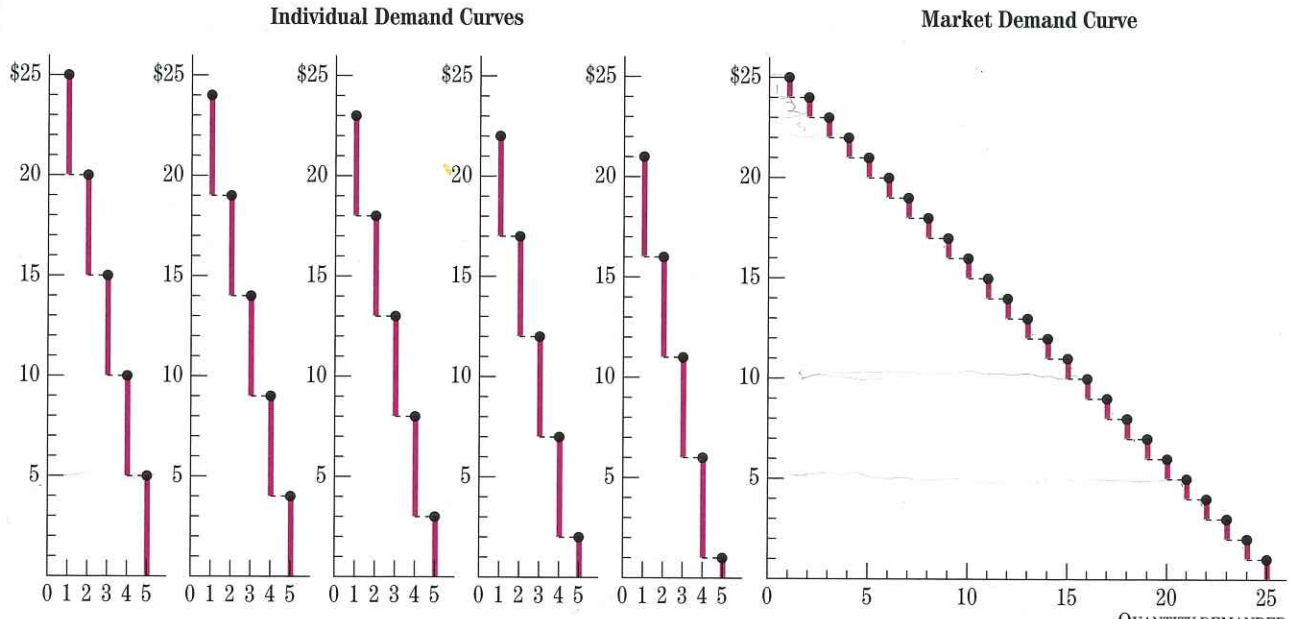
producer surplus: the difference between the price received by a firm for an additional item sold and the marginal cost of the item's production. (Ch. 6)

For example, suppose you have the seller's sheet shown in Table 7.1 and you sell two items. The marginal cost of the first item is \$1, and the marginal cost of the second item is \$6. If you sell the first item for \$15 and the second item for \$10, your gain is (\$15 − \$1) + (\$10 − \$6) = \$18. Observe that each of the terms in this sum is the price minus the marginal cost, or the *producer surplus.*

Predictions of the Competitive Equilibrium Model

Observe that the stage is now set for a market. Buyers are motivated; sellers are motivated. Buyers and sellers can hear or see all the bids and asks. Now what would you predict would happen in this market? Does the outcome depend on the personalities, culture, or intelligence of the buyers and sellers? Does the competitive equilibrium model of consumers (buyers) and firms (sellers) predict the outcome?

■ **Constructing the Model.** Figures 7.3, 7.4, and 7.5 show the competitive equilibrium model that corresponds to this market. The marginal benefits for the buyers and marginal costs for the sellers are used to construct the individual demand curves and the individual supply curves. At the bottom of Figure 7.3 are the marginal benefit sheets of the buyers. Each buyer has one sheet. The graph above each sheet is the individual demand curve derived from the sheet, using the methods of Chapter 5. The market demand curve appears on the right. It is the sum of the individual demand curves. Figure 7.4 shows comparable information for the sellers. There are



Buyer A	
Items	*MB*
1	25
2	20
3	15
4	10
5	5

Buyer B	
Items	*MB*
1	24
2	19
3	14
4	9
5	4

Buyer C	
Items	*MB*
1	23
2	18
3	13
4	8
5	3

Buyer D	
Items	*MB*
1	22
2	17
3	12
4	7
5	2

Buyer E	
Items	*MB*
1	21
2	16
3	11
4	6
5	1

Figure 7.3
Modeling the Demand Side of the Double-Auction Market

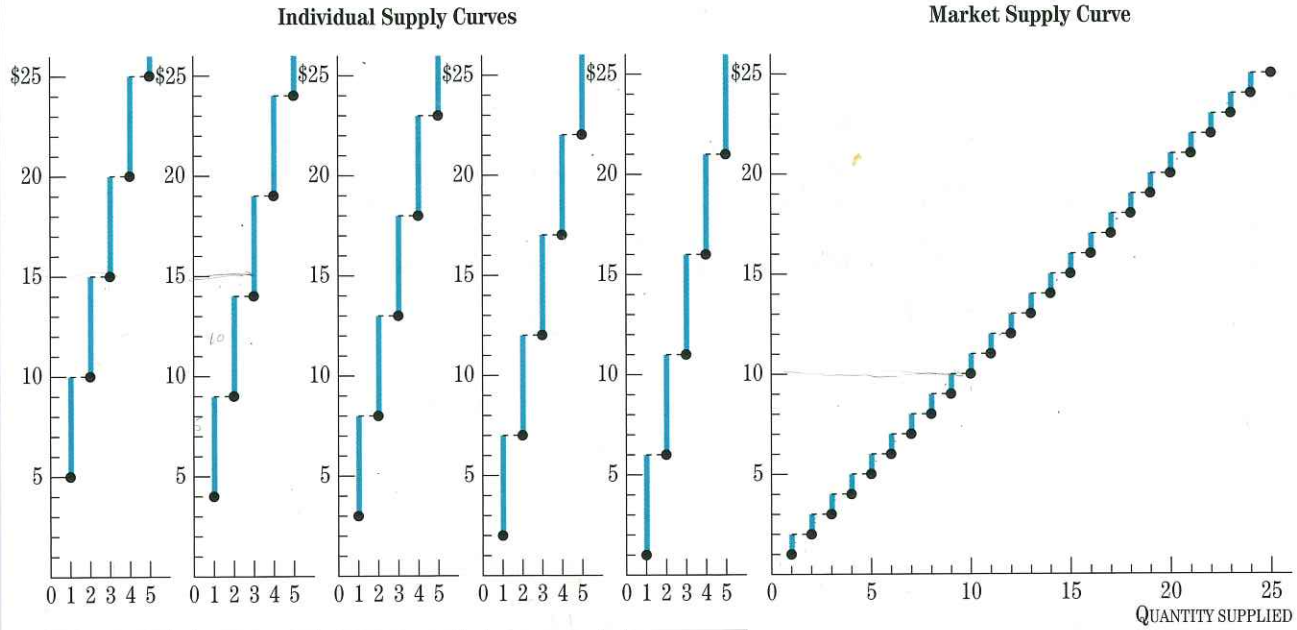
There are five buyers in the market. The marginal benefits (*MB*) for each buyer are shown on that buyer's sheet. The corresponding individual demand curve is above the sheet. The market demand curve is shown on the right. It is the sum of all the items demanded at each price.

marginal costs and upward-sloping individual supply curves. The market supply curve is on the right. It is the sum of the individual supply curves.

Figure 7.5 shows the prediction of the competitive equilibrium model about the price and the quantity that will come out of this market. The demand curve of Figure 7.3 is combined with the supply curve of Figure 7.4. The two curves intersect at a price of \$13 and a quantity of 13 items. In other words, the model predicts that when these 10 people interact in the market, the sellers will sell a total of 13 units, the buyers will buy a total of 13 units, and the market price will be \$13. Is the prediction correct?

■ **Results.** This experimental market has been tried many times, and similar results have occurred each time. First, after one or two trading periods, the price will settle down to about \$13. Sometimes it will be \$12 or \$14, but rarely does it deviate much from \$13. Second, after one or two trading periods, the quantity traded will be very close to 13 units. Again, sometimes it will be slightly more or slightly less.

In other words, the model as shown in Figure 7.5 comes very close to predicting the outcome of the double-auction market. Of course, if you gave different marginal



Seller A	
Items	*MC*
1	5
2	10
3	15
4	20
5	25

Seller B	
Items	*MC*
1	4
2	9
3	14
4	19
5	24

Seller C	
Items	*MC*
1	3
2	8
3	13
4	18
5	23

Seller D	
Items	*MC*
1	2
2	7
3	12
4	17
5	22

Seller E	
Items	*MC*
1	1
2	6
3	11
4	16
5	21

Figure 7.4
Modeling the Supply Side of the Double-Auction Market

There are five sellers in the market. The marginal costs (*MC*) for each seller are shown on that seller's sheet. The corresponding individual supply curve is shown above each sheet. The market supply curve is shown on the right. It is the sum of all items supplied at each price.

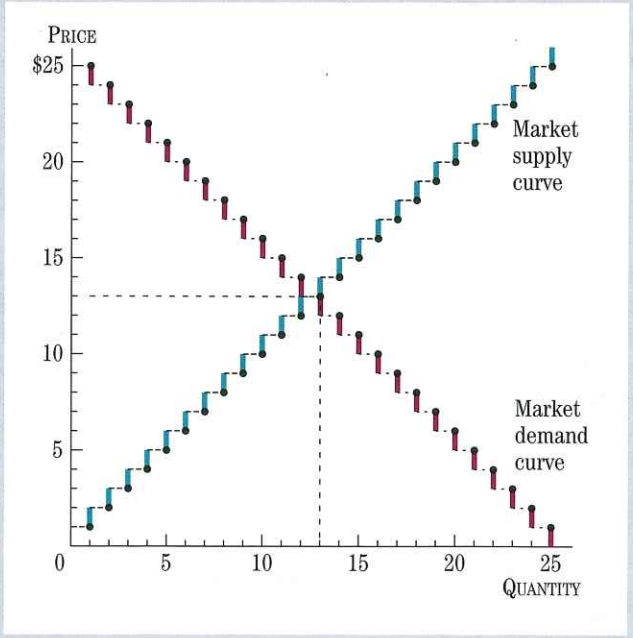


Figure 7.5
Predicted Price and Quantity in the Double-Auction Market
The competitive equilibrium model predicts that the price and quantity traded will be at the intersection of the market supply and demand curves, as shown in the figure. The market demand curve is from Figure 7.3, and the market supply curve is from Figure 7.4. The predicted market price is $13, and the predicted quantity traded is 13 units.

benefits and marginal costs to the buyers and sellers, you would have different supply and demand curves and therefore would get different answers, but the answers would be very close to those predicted by the model. Even though no one individual sets the price—the buyers and sellers are calling out prices in intense competition—the price settles down to the price predicted by the model.

Experimental economists such as Vernon Smith at the University of Arizona and Charles Plott of the California Institute of Technology have found that the model works surprisingly well in predicting the outcome of these experiments, even with a very small number of participants. The finding that the model predicts so well for a very small number of buyers and sellers has been called a "scientific mystery" by Vernon Smith.[1]

Surprise or not, the double-auction market demonstrates both how a market works and how a model works at predicting the outcome of the market. In the next few sections of this chapter, we use this model to measure the gains from trading in the market system. That the model works well in such experiments gives us more confidence in using it for these purposes.

To be sure, even though experiments are set up to mimic the operation of actual markets, they are still done in a laboratory setting. Market participants in the real world may be more or less sophisticated than those in the experiments, and the rules—if there are any—may be more complex in the real world. Thus, even though the experimental confirmation of the model is reassuring, we must remember that it is a model, not reality.

REVIEW

- Experimental markets can be used to test economic models and to demonstrate how markets work.
- The double-auction market is a good type of market for testing the predictions of the competitive equilibrium model.
- The competitive equilibrium model works remarkably well. The price and quantity sold in the double-auction market are usually very close to those predicted by the model.

Are Competitive Markets Efficient?

We have shown how to use the competitive equilibrium model to explain *how* a market works. Now let's use the competitive equilibrium model to see *how well* the market works. Are the quantities produced and consumed in the market efficient?

1. Vernon L. Smith, "Microeconomic Systems as an Experimental Science," *American Economic Review*, Vol. 72, 1982.

The Meaning of Efficient

In general, an inefficient outcome is one that wastes scarce resources, and an efficient outcome is one that does not waste scarce resources. Extremely inefficient economic outcomes are easy to spot. Constructing 275 million new video rental stores each year in the United States (or approximately one store per person) would obviously be wasteful. The workers building the new stores could be building other things that people wanted. If the U.S. economy produced such an outcome, everyone would say it was inefficient; shifting production to fewer video rental stores would clearly make many people better off.

An equally inefficient situation would occur if only one new video rental store a year was built; at that rate it would take more than 2,000 years to build the number of Blockbuster video rental stores that now exist in the United States. In such a situation, shifting production toward more video rental stores would clearly make many people better off.

Both these situations are inefficient because a change in production could make people better off. We might, therefore, define an efficient outcome as one that is so good that there is no change that would make people better off.

The Need for a More Precise Definition. However, because the economy consists of many different people, we need to be more careful in defining efficiency. For every economic outcome, it is possible to make someone better off at the expense of someone else. If someone takes a long-stemmed rose from Maria and gives it to Ken, then Ken is better off but Maria is worse off. More generally, the possibility of transferring a good from one person to another, thereby making someone better off at the expense of someone else, is not an indication that an economic situation is inefficient or wasteful.

However, if there were a situation in which it was possible to change consumption or production in a way that would make someone better off without hurting someone else, then that situation would be inefficient. In such a situation, resources are being wasted, because someone, perhaps many people, could have a better life without someone else being harmed.

Pareto efficient: a situation in which it is not possible to make someone better off without making someone else worse off.

Based on such considerations, economists have developed the following definition of efficiency: An outcome is **Pareto efficient** if it is not possible to make someone better off without hurting someone else. Economists use the term *Pareto* to distinguish this definition of efficiency from other meanings, but the word *efficient* by itself is used when the meaning is clear from the context. Italian economist Vilfredo Pareto is the person who developed this concept of efficiency. Unless we say otherwise, when we use the term *efficient* in this chapter, we mean efficient in the sense of Pareto. If a market is not efficient in the sense of Pareto, then there is something wrong with the market.

Three Conditions for Efficient Outcomes. There are three conditions that must hold if a market outcome is to be efficient in the sense of Pareto efficient.

First efficiency condition: *MB* = *MC* for last item produced.

First, *the marginal benefit* (MB) *must equal the marginal cost* (MC) *of the last item produced.* Why is this condition needed for efficiency? Suppose it did not hold. If the marginal cost is greater than the marginal benefit, then too much is being produced. In the example of producing 275 million video rental stores a year, the marginal cost of producing the 275 millionth video rental store is much greater than the marginal benefit. Reducing production (by a lot) would be appropriate. If the marginal cost is less than the marginal benefit of the product, then too little is being produced. In the example of producing only one video rental store a year, the marginal cost is certainly much less than the marginal benefit; more production would be appropriate. Only when marginal benefit is equal to marginal cost is the economic outcome efficient. This must occur for all goods from video rental stores to roses.

One way to better appreciate this condition is to imagine that you grow your own roses in your own garden. Clearly you would never produce more roses if the marginal cost to you was greater than the marginal benefit to you. But you would produce more roses if your marginal benefit from more roses was greater than your marginal cost. Only when marginal benefit equals marginal cost would you stop producing and consuming more.

Second efficiency condition: Every producer's *MC* is the same.

The second condition for efficiency relates to the production of goods at different firms. It is that *the marginal cost of a good should be equal for every producer.* Again, if this were not the case, then production could be increased without cost. For example, if Hugo's rose garden could produce an extra dozen roses at a marginal cost of $10 and Mimi's rose garden could produce an extra dozen roses at a marginal cost of $50, then it would make sense for Hugo's garden to increase production and for Mimi's garden to decrease production. Mimi could take the $50 she saved by producing less and have more than enough to pay Hugo's costs of producing an extra dozen. Only when the marginal costs are the same is there no way to increase production without cost. Note that it is not necessary for Hugo and Mimi or any other producer to be the same or even to have the same total costs; all that we require for efficiency is that the *marginal* costs be the same.

Third efficiency condition: Every consumer's *MB* is the same.

The third condition for efficiency relates to the allocation of goods to different consumers. It is that *the marginal benefit of consuming the same good should be equal for all consumers.* If the marginal benefits were not equal, then there could be a gain for some people with no loss for anyone else. For example, suppose Ken's marginal benefit from roses was $3 and Maria's was $1; then if Maria sold roses to Ken for $2, both would be better off. But if their marginal benefits were the same, then no improvement for one without harming the other would be possible.

In sum, there are three conditions for efficiency: (1) the marginal benefit equals the marginal cost for the last item produced; (2) the marginal cost of producing each good is equal for all producers; and (3) the marginal benefit from consuming each good is equal for all consumers.

Is the Market Efficient?

Given the three conditions for efficiency, can we say that the market is efficient? The competitive equilibrium model provides us with a quick answer to that question.

Here's the reason in a nutshell why the first condition is satisfied:

At a market equilibrium:

$$P = MB \text{ and } P = MC$$

Thus since $P = P$, we must have

$$MC = MB$$

According to the model of consumer behavior in Chapter 5, an individual consumer chooses a quantity of a good such that *price equals marginal benefit*—that is, $P = MB$. This equality holds for every consumer at every point on the market demand curve. Remember that the marginal benefit is the willingness to pay dollars to consume an additional amount of a good. According to the model of firm behavior in Chapter 6, a firm produces a quantity of a good such that *price equals marginal cost.* That is, $P = MC$. This equality holds for every firm at every point on the market supply curve. At a point of intersection of the supply curve and the demand curve, both of these conditions must hold because the point of intersection is on both the supply curve and the demand curve and the price P is the same. That is, $P = MB$ and $P = MC$ simultaneously. This implies that at the quantity produced by the market, *marginal benefit equals marginal cost.* That is, $MB = MC$. This is true of every good.

Thus we have proved that a competitive market satisfies the first condition of efficiency. The marginal cost of producing roses, grapes, bread, peanuts, or automobiles is equal to the marginal benefit that people get from consuming them. This occurs without any person coordinating consumers and producers. Producing more or less of the item will only lead to a violation of this key equality between marginal cost and marginal benefit.

Here's the reason in a nutshell why the second condition is satisfied:

Hugo's $MC = P$
Mimi's $MC = P$

Thus

Hugo's MC = Mimi's MC

Here's the reason in a nutshell why the third condition is satisfied:

Maria's $MB = P$ and
Ken's $MB = P$

Thus

Maria's MB = Ken's MB

first theorem of welfare economics: the conclusion that a competitive market results in an efficient outcome; sometimes called the "invisible hand theorem"; the definition of efficiency used in the theorem is Pareto efficiency.

To better appreciate the result, again imagine that you grow your own roses in your own garden. Clearly, you would never grow more roses if the marginal benefit to you was less than your marginal cost. What is striking is that when you do not grow your own roses or even when you do not know anything about growing roses, the marginal benefit of more roses to you will be equal to the marginal cost of producing more roses.

The result is illustrated in Figure 7.6. At the market equilibrium quantity (point *E*), the marginal cost (the point on the supply curve) is equal to the marginal benefit (the point on the demand curve). At any other point, either marginal benefit will be greater than marginal cost or marginal benefit will be less than marginal cost.

The other two criteria for efficiency also hold in a competitive market. To see this, it will help to look back at Figure 7.2. Observe that in a market equilibrium, the marginal cost for the producers is the same, because *they all face the same price;* along each of their individual supply curves, all producers—Hugo, Mimi, and others—set marginal cost equal to the price. Similarly, in a market equilibrium, *all consumers*—Maria, Ken, and others—*face the same market price.* Hence, their marginal benefits are all equal, because on each of their individual demand curves the marginal benefit equals the price. Thus, it is not possible to make one person better off without hurting someone else. In a competitive market, the marginal benefits are equal. Thus, there is no improvement for one that does not hurt someone else.

In sum, for each good produced in a competitive market, (1) the marginal benefit equals the marginal cost of the last item produced, (2) the marginal cost is equal for all producers in the market, and (3) the marginal benefit is equal for all consumers in the market. Thus, we can say that the competitive market is Pareto efficient.

The proposition that competitive markets are efficient is one of the most important in economics, so much so that when it is proven with the mathematics necessary to keep track of many different goods and time periods, it is called the **first theorem of welfare economics.** The word *theorem* reflects the mathematics used in the advanced proof of the idea. The word *welfare* means that the theorem is

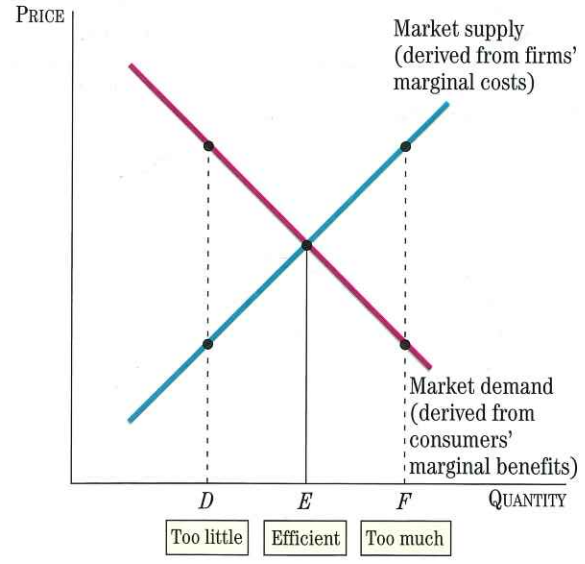


Figure 7.6
The Efficiency of the Market: Marginal Benefit Equals Marginal Cost
Only at quantity *E* is the marginal benefit of an extra unit equal to the marginal cost of an extra unit. Point *D* is not efficient because the marginal benefit of an extra unit is greater than the marginal cost of producing it. Part *F* is also not efficient because the marginal cost of producing an extra unit is greater than the marginal benefit.

READING THE NEWS ABOUT... Ticket Scalping – or Taking Control of Ticket Sales

Ticket scalping—selling a ticket to an event at a higher price than its face value—is a fascinating subject, illustrating how the workings of free markets intertwine with the complexity of human emotion. A common view of scalping is expressed in this comment written in an MIT newspaper by a student who had just bought an extra ticket to his own graduation from a scalper: "There is no disputing the fact that we live in a capitalist society where the laws of supply and demand apply to many transactions, but are we not an academic community of concerned individuals? Graduation tickets were not meant to be sold to the highest bidder, like the scalpers who peddle tickets to the latest rock concert or playoff sporting event." (*The Tech*, 5/27/94, p 4.)

Musicians (and possibly some sports teams) might agree more with the MIT student than he suspects. The Internet has generated an exceedingly efficient way of supplying tickets to people who demand them. But, as described in the article below, many musical acts view this type of sale as detrimental to their livelihood in the long run. Losing control of the sale may prevent a type of interaction (between fan and musical performer) that many bands would prefer to encourage. Why do you think bands want to get back control of their tickets sales and remove the middleman from this economic interaction?

A Guide to Tickets
Where Have All the Good Seats Gone?

by Steve Morse, *Boston Globe* Staff
May 1, 2005

Tired of scalpers marking up ticket prices and frustrated with high service fees from brokers, many music acts are taking more control of their ticketing as a way to connect with fans.

The biggest trend is for bands to offer tickets through their online fan club sites, with some charging annual membership fees. U2 charges $40, the Dave Matthews Band charges $35, and Madonna charges about $38.

The intent is to get tickets into the hands of diehard fans before the tickets go on sale to everyone else. Typically, bands hold Internet "presales" to club members a week or two before a public sale, which is usually handled by broker Ticketmaster.

But for fans, the clubs certainly don't guarantee tickets. Nor cheap tickets either. Last year Madonna's fan site offered $700 VIP seats. But bands for the most part offer lower service fees and throw in discounts on merchandise and other sweeteners to promote artist-to-fan interaction without a middleman.

One concern is that some scalpers just join the fan clubs, which was a problem with a recent U2.com presale. But bands are making vigorous attempts to cross-check names, credit card numbers, and e-mail and postal addresses to weed out violators who may be buying tickets in bulk, as well as work with online auction site eBay to identify culprits.

"You can't always make everyone happy, but we do the best job we can," says Coran Capshaw, who manages the Dave Matthews Band and heads Musictoday.com, which handles fan sites for Madonna, Kenny Chesney, Britney Spears, and other top-name acts.

Musictoday.com operates out of Charlottesville, Va., and employs 200 people who coordinate the ticketing and the discounted merchandise and other privileges (from CD samplers to special magazines) that come with the memberships. Tens of thousands of fans have joined

about the overall well-being of people in the economy (the word *welfare* is synonymous with "well-being," not with a transfer payment to a poor person). The word *first* is used to distinguish this theorem from the second theorem of welfare economics, which states the converse of the first: Any Pareto efficient outcome can be obtained via a competitive market.

Efficiency and Income Inequality

income inequality: disparity in levels of income among individuals in the economy.

Efficiency is a very important goal of an economic system, but it is not the only goal. Another goal is that no one, or at least as few people as possible, falls into dire economic circumstances. For example, reducing **income inequality** to an amount that makes poverty a rare occurrence is also a desirable goal in most economic systems.

the Matthews band's fan site called "Warehouse," says Capshaw, who declined to be more specific.

Capshaw says his role model was the Grateful Dead, who pioneered direct-mail marketing. The Dead would take up to 50 percent of a venue's seats and distribute customized souvenir tickets.

Most bands can now only get 8 to 10 percent of a show's tickets because they run up against tight contracts between Ticketmaster and the venues. In the touring business, concert sites such as the Tweeter Center and Bank of America Pavilion make agreements with Ticketmaster to sell tickets and charge customers a fee for that service. Ticketmaster then shares part of the service fees with the music promoter and the venue, but not the bands.

Acts that have lobbied to control their tickets for years, such as the Dave Matthews Band and String Cheese Incident, can get a higher percentage of a show's tickets. When Ticketmaster tried to cut back String Cheese's allotment in 2003, the band filed an antitrust suit against the ticketing agency. The suit was settled out of court for undisclosed terms, and String Cheese now says it is pleased with its current share.

John Pleasants, chief executive of Ticketmaster, said ticket presales have gained in popularity the last five years, in part because bands are trying to develop new revenue streams through their fan clubs. Pleasants said Ticketmaster runs the presales for many bands.

"Artists are the people who are at the top of the value chain," Pleasants said. "These are the people who create the content and there is no question that, as a generalization, these folks are looking to get more compensation for what they do, as opposed to less, and so they will look for opportunities like fan clubs or extra merchandise sales."

Some venue promoters dislike fan club sites. A Clear Channel Entertainment executive says a fan club presale can hurt the subsequent public sale of tickets because many nonmembers will assume the best seats are already taken. For that reason, Clear Channel in Boston, the executive adds, makes sure that many good tickets are still held for the public sale.

Lower service fees are another key reason fan club sites have flourished. The jam band String Cheese Incident tacks on an average fee of $4 per ticket, but that's often less than half of what Ticketmaster might charge.

"Hopefully by lowering the fees, the fans can afford to see more music," said String Cheese manager Mike Luba.

His group first started selling tickets out of a candle shop run by bassist Keith Moseley's brother, Kevin, a decade ago in Durango, Colo. That evolved into SCI Ticketing and is now called Baselineticketing.com, which has since signed up more than 20 acts, from punk superstars Green Day to Boston breakout band the Dresden Dolls. Most of Baselineticketing.com's acts do not charge membership fees because they are opposed to a "two-tiered process" that allows wealthier fans to have an advantage over those of more modest means.

"Still, you might rather pay $40 for a membership than pay scalpers for tickets," said Bob Grossweiner, a senior editor with liveDaily.com, an online magazine about the music industry.

Dashed expectations can anger fans, as U2 learned when the presale for the first leg of its American tour went awry. Many members didn't get tickets, fans complained, and the band apologized.

"We weren't expecting the amount of members who joined," said Sebastian Clayton, brother of U2 bassist Adam Clayton, who runs the site. He said the system was overwhelmed when tickets became available in January.

"We were left in the lurch and weren't able to fill everyone's order," said Sebastian. He said the problem was fixed after U2 tacked on another leg to its tour.

Fan club sites can be big business. Take U2.com. With 100,000 members paying a $40 annual fee, that amounts to $4 million of revenue.

"It's capitalists outfoxing other capitalists," Grossweiner says.

It is important to emphasize that efficiency and income equality are not the same thing. An allocation of bread between Hugo and Mimi is efficient if their marginal benefit of bread is the same and if the marginal benefits equal the marginal cost of bread. Then there is no mutually advantageous trade of bread between Hugo and Mimi that will make one better off without hurting the other.

However, suppose that Hugo has a low income, earning only $7,000 per year, and that Mimi has a high income, earning $70,000 per year. Suppose a severe drought raises the price of wheat and thus the price of bread. If the price of bread in the market gets very high, say, $3 a loaf, then Hugo will be able to buy few loaves of bread and may go hungry, especially if he has a family. In this case, the economy gets good marks on efficiency grounds but fails miserably on income inequality grounds.

To remedy the situation, a common suggestion is to put price controls on bread. For example, to help Hugo and others like him, a law might be passed requiring that

bread prices not exceed $.50 a loaf. Although this may help the income inequality problem, it will cause great inefficiency because it interferes with the market. At $.50 a loaf, bread producers will not produce very much, and Mimi will probably start buying bread to feed the birds, wasting scarce resources.

A better solution to the income inequality problem is to transfer income to Hugo and other low-income people from Mimi and other high-income people. With a transfer of income—say, through a tax and an income-support payment to the poor—the market would be able to function and the gross inefficiencies caused by price controls on bread would not occur. Even at the high price of bread, Hugo will be able to eat, perhaps buying some rice or a bread substitute, and the bread, which is so expensive to produce, will not be wasted on the birds. We will see that such transfers have advantages and disadvantages. Compared to price controls, their main advantage is that they allow the market to operate efficiently.

The temptation to deal with income inequality problems in ways that interfere with the efficiency of the market is great in all societies. Price ceilings (rent controls) on rental apartments in some U.S. cities, which we examined in Chapter 3, are one example. But this interference wastes economic resources.

REVIEW

- Economic inefficiency implies a waste of resources. A Pareto efficient outcome is one in which no person's situation can be improved without hurting someone else. A key criterion for Pareto efficiency is that production and consumption be such that the marginal benefit of a good equals its marginal cost.
- One of the most desirable features of competitive markets is that at the equilibrium level of production, marginal benefit equals marginal cost. On the demand curve, marginal benefit equals the market price. On the supply curve, marginal cost equals the market price. Because the quantity demanded equals the quantity supplied in equilibrium, together these imply that marginal benefit equals marginal cost.
- Thus competitive markets are efficient. Any change in consumption or production that makes one person better off must make someone else worse off.
- Efficiency is not the same thing as income equality. An efficient outcome can coexist with an unequal outcome.

Measuring Waste from Inefficiency

We know from Chapters 5 and 6 that consumer surplus and producer surplus are measures of how much consumers and producers gain from buying and selling in a market. The larger these two surpluses are, the better off people are.

Maximizing the Sum of Producer Plus Consumer Surplus

An attractive feature of competitive markets is that they maximize the sum of consumer and producer surplus. Producer and consumer surplus are shown in the market supply and market demand diagram in Figure 7.7. Recall that the producer

Another way to think about the lightly shaded areas in the graphs: The sum of consumer surplus plus producer surplus is the triangular area between the demand curve and the supply curve—shown by the lightly shaded area in the middle graph of Figure 7.7. The graph shows another way to think about this sum: The sum of consumer surplus plus producer surplus equals the marginal benefit minus the marginal cost of all the items produced.

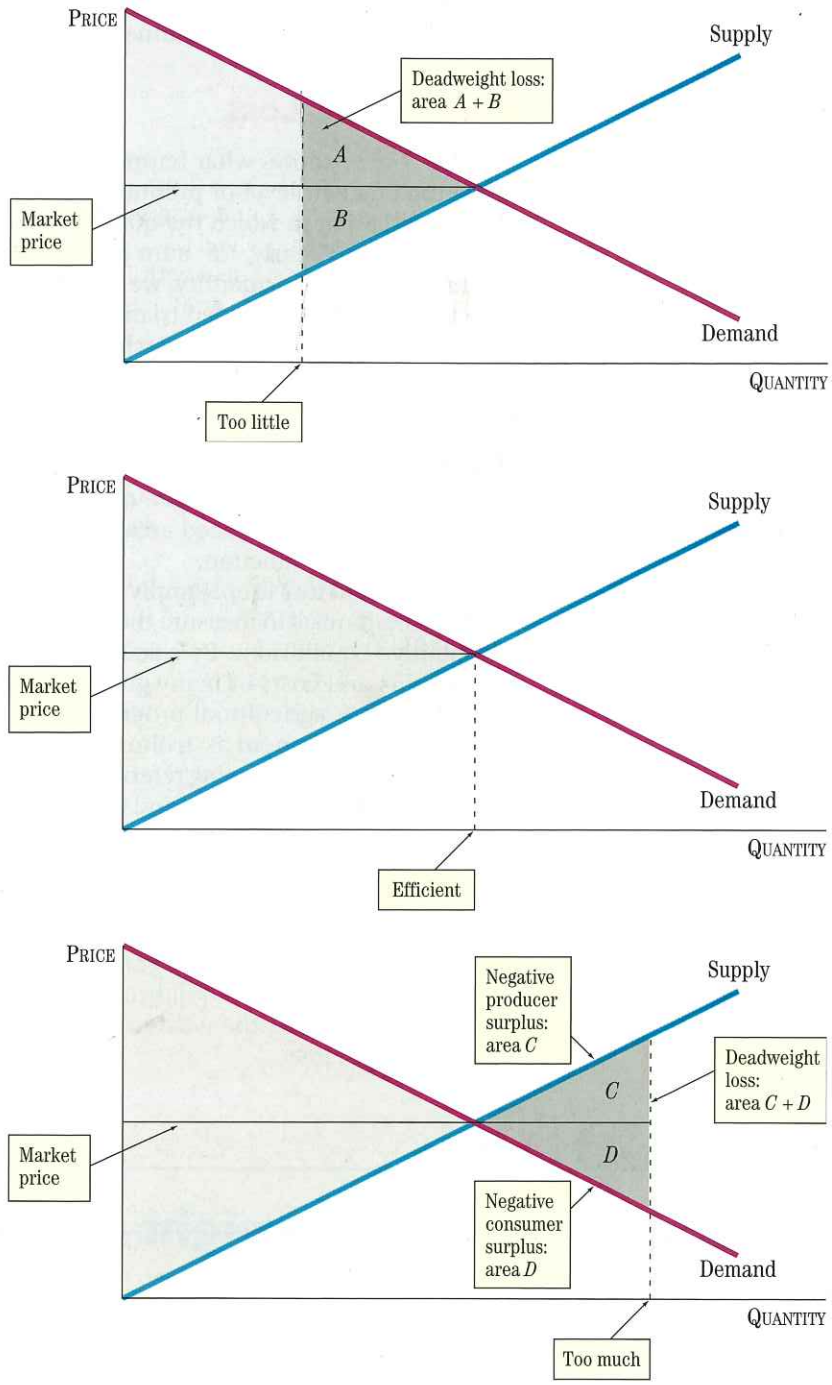


Figure 7.7
Measuring Economic Loss
When production is less or more than the market equilibrium amount, the economic loss is measured by the loss of consumer surplus plus producer surplus. In the top diagram, the quantity produced is too small. In the bottom diagram, it is too large. In the middle diagram, it is efficient.

surplus for all producers is the area above the supply curve and below the market price line. The consumer surplus for all consumers is the area below the demand curve and above the market price line. Both the consumer surplus and the producer surplus are shown in Figure 7.7. The lightly shaded gray area is the sum of consumer surplus plus producer surplus. The equilibrium quantity is at the intersection of the two curves. At this point, consumer surplus plus producer surplus is maximized.

Deadweight Loss

Figure 7.7 also shows what happens to consumer surplus plus producer surplus when the efficient level of production does not occur. The top panel of Figure 7.7 shows a situation in which the quantity produced is lower than the market equilibrium quantity. Clearly, the sum of consumer and producer surplus is lower. By producing a smaller quantity, we lose the amount of the consumer and producer surplus in the darkly shaded triangular area $A + B$. The bottom panel of Figure 7.7 shows the opposite situation, in which the quantity produced is too high. In this case, we have to subtract the triangular area $C + D$ from the lightly shaded area on the left because price is greater than marginal benefit and lower than marginal cost, which means that consumer surplus and producer surplus are negative in the area $C + D$. In both the top and bottom panels of the figure, these darkly shaded triangles are a loss to society from producing more or less than the efficient amount. Economists call the loss in this darkly shaded area the **deadweight loss.** It is a measure of the waste from inefficient production.

deadweight loss: the loss in producer and consumer surplus due to an inefficient level of production.

Deadweight loss is not simply a theoretical curiosity with a gruesome name; it is used by economists to measure the size of the waste to society of deviations from the competitive equilibrium. By calculating deadweight loss, economists can estimate the benefits and costs of many government programs. When you hear or read that the cost of U.S. agricultural programs is billions of dollars or that the benefit of a world-trade agreement is trillions of dollars, it is the increase or decrease in deadweight loss that is being referred to. In order to compute the deadweight loss, all we need is the demand curve and the supply curve.

REVIEW

- Competitive markets maximize producer surplus plus consumer surplus.
- If the quantity produced is either greater or less than the market equilibrium amount, the sum of consumer surplus plus producer surplus is less than at the market equilibrium. The decline in consumer plus producer surplus measures the waste from producing the wrong amount. It is called deadweight loss.

CASE STUDY

Price Controls and Deadweight Loss in the Milk Industry

Since the 1930s, the federal government has intervened in the milk market (and other agricultural markets) in order to stabilize farm prices and provide some income protection for U.S. farmers. The government has used a combination of complex regulations that include government purchases and subsequent disposal of dairy products, import restrictions, export subsidies, and pricing mechanisms depending on the

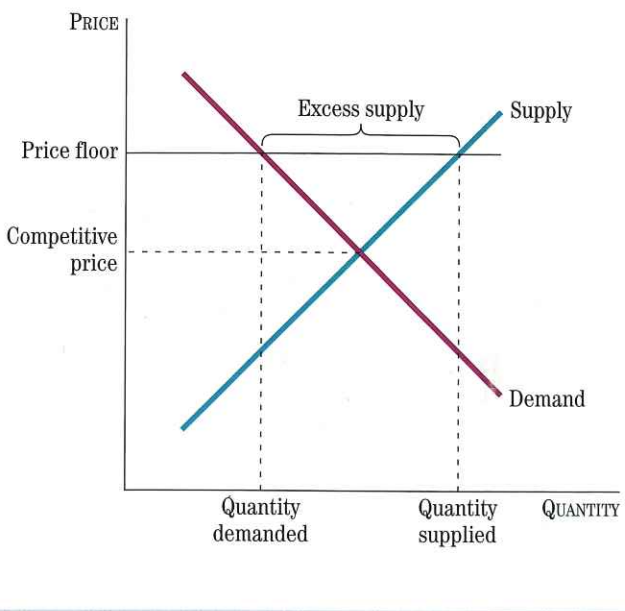


Figure 7.8
Price and Quantity Effects of a Price Floor
If the price floor is set higher than the competitive market price, the quantity demanded by consumers decreases and the quantity supplied by firms increases, creating excess supply.

location and purpose of the production of milk. We can see how price controls lead to deadweight loss by looking more closely at one of these programs.

The Food and Agriculture Act of 1977 was aimed at sustaining higher prices received by dairy farmers. Figure 7.8 shows a stylized representation of the milk market with a price floor. As we know, the competitive market price occurs when the quantity demanded equals the quantity supplied, but the higher price floor mandated by the government reduced the quantity demanded and gave farmers an incentive to produce more milk, causing excess supply. To support the price floor, the government purchased the excess supply of milk in the form of dry milk, butter, and cheese. Of course, there was a cost to this program: close to $2 billion a year in net government expenditures in the early 1980s. In the late 1980s and early 1990s, in an effort to reduce the excess supply of milk while keeping prices high, the federal government taxed farmers who increased their milk production.

Figure 7.9 shows the reduction in consumer surplus, the increase in producer surplus, and the excess supply of milk that is purchased by the government. In 1994 economists Peter Helmberger and Yu-Hui Chen estimated what would happen if the government deregulated the milk market. In the short run, they found that consumer surplus would increase by $3.9 billion a year, producer surplus would decrease by $4 billion, and net government expenditures would decrease by $600 million, eliminating a deadweight loss of $500 million a year. As you can see, the price floor is more expensive than directly transferring money from consumers to farmers, as explained in the Efficiency and Income Inequality section.

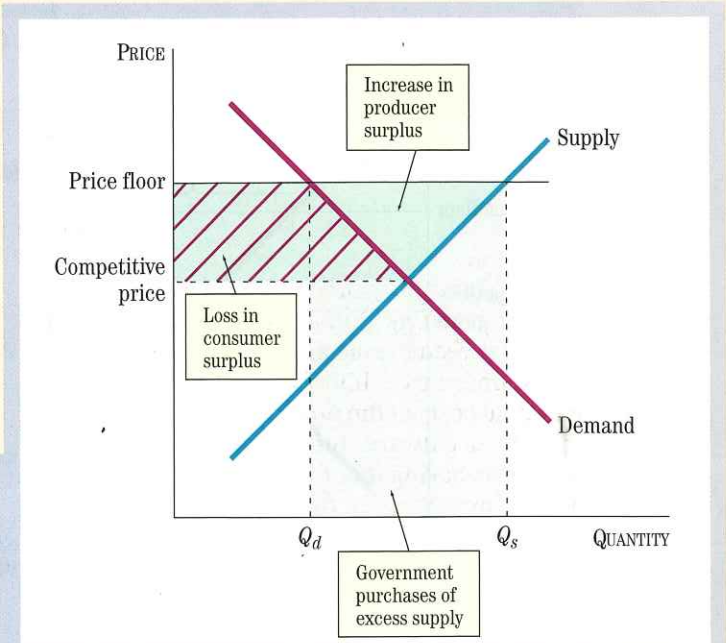


Figure 7.9
A Costly Price Floor Program
Each unit of excess supply has to be purchased by the government at a price higher than the competitive market price, making this a very costly program.

The Federal Agricultural Improvement and Reform (FAIR) Act of 1996 mandated the elimination of the price support program by the end of 1999. However, the dairy subsidies were soon reinstated by the farm bill signed by President Bush in May 2002, which increased total agricultural subsidies from $100 billion to close to $200 billion a year. The current system of dairy subsidies chose the market price of drinking milk in Boston as the standard for the rest of the country. When that price falls below $16.94 per hundred pounds, all U.S. dairy farmers receive a governmental subsidy of 45 percent of the difference between the Boston market price and $16.94. What do you think will be the effects of this legislation on the milk market?

REVIEW

- Price floors hurt consumers and usually benefit producers. The net effect is a loss to society.
- Agricultural subsidies in the United States were doubled in 2002.

The Deadweight Loss from Taxation

Another important application of deadweight loss is in estimating the impact of a tax. To see how, let's examine the impact of a tax on a commodity like gasoline. We will see that the tax shifts the supply curve, leads to a reduction in the quantity produced, and reduces the sum of producer surplus plus consumer surplus.

A Tax Paid by a Producer Shifts the Supply Curve

A tax on sales is a payment that must be made to the government by the seller of a product. The tax may be a percentage of the dollar value spent on the products sold,

in which case it is called an *ad valorem tax*. A 6 percent state tax on retail purchases is an ad valorem tax. Or it may be proportional to the number of items sold, in which case the tax is called a *specific tax*. A tax on gasoline of $.50 per gallon is an example of a specific tax.

Because the tax payment is made by the producer or the seller to the government, the immediate impact of the tax is to add to the marginal cost of producing the product. Hence, the immediate impact of the tax will be to shift the supply curve. For example, suppose each producer of gasoline has to send a certain amount, say, $.50 per gallon produced and sold, to the government. Then $.50 must be added to the marginal cost per gallon for each producer.

The resulting shift of the supply curve is shown in Figure 7.10. The vertical distance between the old and the new supply curves is the size of the sales tax in dollars. The supply curve shifts up by this amount because this is how much is added to the marginal costs of the producer. (Observe that this upward shift can just as accurately be called a leftward shift because the new supply curve is above and to the left of the old curve. Saying that the supply curve shifts up may seem confusing because when we say "up," we seem to be meaning "more supply." But the "up" is along the vertical axis, which has the price on it. The upward, or leftward, movement of the supply curve is in the direction of less supply, not more supply.)

A New Equilibrium Price and Quantity

What does the competitive equilibrium model imply about the change in the price and the quantity produced? Observe that there is a new intersection of the supply curve and the demand curve. Thus, the price rises to a new, higher level, and the quantity produced declines.

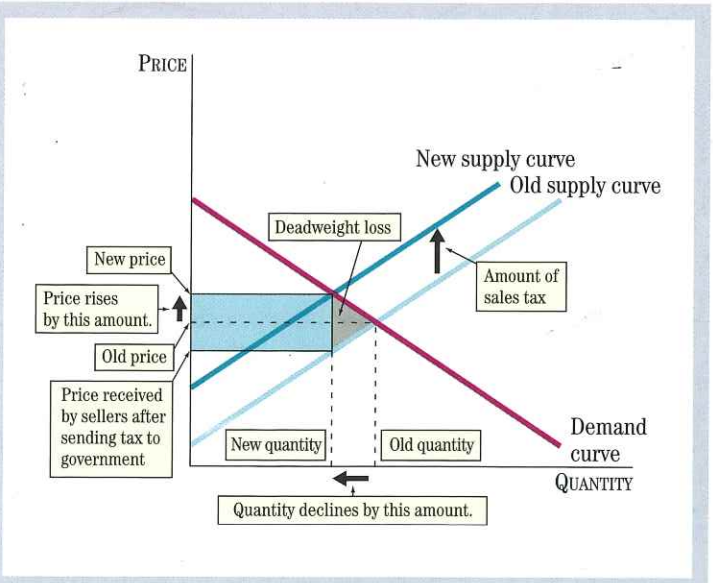


Figure 7.10
Deadweight Loss from a Tax
In this graph the dark triangle represents the deadweight loss and the blue rectangle the amount of tax revenue that goes to the government. The sales tax, which is collected and paid to the government by the seller, adds to the marginal cost of each item the producer sells. Hence, the supply curve shifts up. The price rises, but by less than the tax increase.

The price increase, as shown in Figure 7.10, is not as large as the increase in the tax. The vertical distance between the old and the new supply curves is the amount of the tax, but the price increases by less than this distance. Thus the producers are not able to "pass on" the entire tax to the consumers in the form of higher prices. If the tax increase is $.50, then the price increase is less than $.50, perhaps $.40. The producers have been forced by the market—by the movement along the demand curve—to reduce their production, and by doing so they have absorbed some of the tax increase.

Deadweight Loss and Tax Revenue

Now consider what happens to consumer surplus and producer surplus with the sales tax. Because the total quantity produced is lower, there is a loss in consumer surplus and producer surplus. The right part of the triangle of consumer plus producer surplus has been cut off, and this is a measure of the deadweight loss to society, as shown in Figure 7.10. This loss occurs despite the fact that the tax revenue going to the government is used for financing government activity. The deadweight loss is incurred because there is a movement of production away from the efficient level. Taxes may be necessary to

Economics in Action

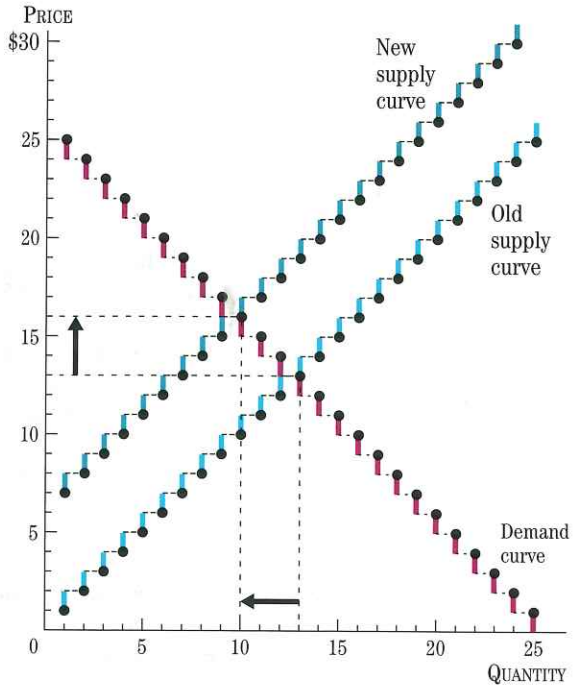
Tax Prediction Passes Economists' Test

The competitive equilibrium model gives a remarkably precise prediction that an increase in a tax on gasoline will increase the price of gasoline, but by no more than the tax increase. This model is also very useful because it predicts that gasoline stations will pass on some but not all of the tax to consumers. The same is true for any other good. If there is a tax on long-stemmed roses, producers such as Hugo and Mimi will have to send the tax payment to the government, but they will find that the price they get for roses will increase as a partial offset. Consumers such as Ken and Maria will have to pay more for roses, but their price increase will be less than the amount of the tax.

But how accurate is the prediction? Economists have tested the prediction with experimental double-auction markets and found that it passes the test. Here is how the test works: Add a fixed amount—say, $6—to the marginal cost sheets of sellers in the double auction. Replacing the original sellers' sheets in Figure 7.4 with new sheets in which marginal cost is $6 higher would show the impact of a $6 tax.

The effect of such a change on the market supply curve is shown on the graph. It shifts up by exactly $6. According to the competitive equilibrium model, the price should rise by $3 and production should fall by 3 units. The price increase should be less than the tax increase, with about half the increase passed on to the buyers, according to the model.

Economists who run such experimental markets find that in this case, the price rises to about $16 and the quantity falls to 10 units, compared with $13 and 13 units without the tax, much as predicted by the model. This occurs even if the buyers know nothing about the tax. The interaction of only a few sellers and buyers in the market with very limited information results in the outcome predicted by the competitive equilibrium model.



finance the government, but they cause a deadweight loss to society. In Chapter 14, where we study the effects of different types of taxes, we will show that the deadweight loss depends on the price elasticity of supply and demand.

Figure 7.10 also tells us how much tax revenue goes to the government in the case of a specific tax. The tax revenue is the tax times the number of items sold. If the tax is $1 and 100 items are sold, the tax revenue is $100. This amount is shown by the blue rectangle. Some of what was producer surplus and consumer surplus thus goes to the government. Another portion, the deadweight loss, is no longer available. No one gets it.

REVIEW

- The impact of a tax on the economy can be analyzed using consumer surplus and producer surplus.
- Taxes are necessary to finance government expenditures, but they lower the production of the item being taxed.

- The loss to society from the decline in production is measured by the reduction in consumer surplus and producer surplus, the deadweight loss due to the tax.

Informational Efficiency

We have shown that a competitive market works well in that the outcome is Pareto efficient. For every good, the sum of consumer surplus and producer surplus is maximized. These are important and attractive characteristics of a competitive market.

Another important and attractive characteristic of a competitive market is that the market processes information very efficiently. For example, in the double-auction market, the price reflects the marginal benefit for every buyer and the marginal cost for every seller. If a government official were asked to set the price in a real market, there would be no way that such information could be obtained, especially with millions of buyers and sellers. In other words, the market seems to be informationally efficient. Pareto efficiency is different from this *informational efficiency*.

In the 1930s and 1940s, as the government of the Soviet Union tried to centrally plan production in the entire economy, economists became more interested in the informational efficiency of markets. One of the most outspoken critics of central planning, and a strong advocate of the market system, was Friedrich Hayek, who emphasized the importance of the informational efficiencies of the market. In Hayek's view, a major disadvantage of central planning—where the government sets all the prices and the quantities—is that it is informationally inefficient.

If you had all the information about all the buyers and sellers in the double-auction market, you could set the price to achieve a Pareto efficient outcome. To see Hayek's point, it is perhaps enough to observe such experimental markets and see that without private information about every buyer and seller, you or any

Coordination Without a Market *Although prices provide a valuable coordination role in a market economy, some activities are better coordinated without the market. It would not be efficient to coordinate each of the hand and foot movements of these 100 skydivers with prices.*

government official would not know where to set the price. Complicate this with millions of buyers, millions of different products, and rapidly changing tastes and technology, and you can quickly comprehend Hayek's arguments. However, economists do not have results as neat as the first theorem of welfare economics to prove Hayek's point. The reason is that in some situations, the market would be unwieldy, and it is difficult to describe these situations with any generality.

Consider the example of coordinating the members of a marching band consisting of several different instruments and several different players. Suppose you were asked how to coordinate the members of a marching band through a price system in a market! You might set a price for playing loud versus soft and then vary the price according to how loud you wanted the band to play. But using the price system to conduct a real band would be an impossible task. It would be better to conduct the marching band without prices and without a market, just as all marching bands in the world are conducted. Coordinating millions of producers and consumers of roses by a central conductor is just as difficult as coordinating the members of a marching band by a price system. Rose production and consumption is handled well by the market and poorly by a central conductor. On the other hand, a marching band is handled poorly by the market and well by a central conductor.

Two obvious difficulties arise in using a market system to coordinate activities like a marching band. First, prices will not bring about a sufficiently precise or speedy response. It is essential that the flute and the sax start playing at the same moment; a one-second delay will turn music into noise. It is better to tell the musicians to play this note at this volume at this time. Second, it is possible for the conductor to get information about each band member. A band leader knows which band member is capable of doing what.

In most situations in which the informational advantage of the price system and the market is not large and in which great precision in coordination is required, organizations spring up. Musicians form a band, a community forms a police force, and so on.

REVIEW

- The market has the ability to process information efficiently. Market experiments demonstrate informational efficiency. The lack of informational efficiency is a key reason why central planning does not work well in complex and changing environments.
- For some activities, however, the market has few informational advantages. A production process in which exact timing is essential will be poorly coordinated through prices. In almost all these situations, firms or organizations form and replace market transactions.

Conclusion

Adam Smith's idea of the "invisible hand" is perhaps the most important discovery in economics: Individuals, by freely pursuing their own interests in a market economy, are led as if by an invisible hand to an outcome that is best overall. The first theorem of welfare economics is the modern statement of Adam Smith's famous principle; in tribute to Smith's seminal idea, we call it the "invisible hand theorem," although the theorem was not actually proved by economists until the mid-twentieth

century. Understanding why, and under what circumstances, the invisible hand theorem is true is an important part of thinking like an economist.

Understanding the theorem has required an investment in economic model building: The behavior of consumers and the behavior of firms were combined into a competitive equilibrium model describing how consumers and firms interact in markets. This model is an embellishment of the supply and demand model we used in Chapters 3 and 4. Experimental markets demonstrate that the model works well in predicting actual outcomes.

Building the competitive equilibrium model has had payoffs beyond understanding this most important theorem in economics. Armed with the ideas of consumer surplus and producer surplus, we can now measure the costs of deviations from the competitive market equilibrium. Such measures are used by economists to assess the costs and benefits of government programs that interfere, for bad or good, with the market outcomes. Starting with Chapter 10, we will see that deviations from the competitive market equilibrium are caused by monopolies and other factors. But first we will look more closely at how costs and production within individual firms and competitive industries change over time. We do this in Chapters 8 and 9.

KEY POINTS

1. The interaction of producers and consumers or buyers and sellers in a market can be explained by the competitive equilibrium model.
2. Processing information and coordinating and motivating millions of consumption and production decisions is difficult, but the market is a device that can do the job remarkably well.
3. The competitive equilibrium model keeps track of the individual decisions of consumers and producers.
4. Even with only a few sellers and buyers, experimental markets appear to be well explained by the competitive equilibrium model.
5. An outcome is Pareto efficient if it is not possible to change production or consumption in a way that will make one person better off without hurting someone else.
6. A competitive market is Pareto efficient.
7. In a competitive market, marginal benefit equals marginal cost for the last item produced, and the sum of producer surplus and consumer surplus is maximized.
8. Deviations from the Pareto efficient outcome create a loss to society called deadweight loss.
9. Deadweight loss is caused by a tax that reduces the quantity produced.
10. The market system is also informationally efficient. However, there are no general theorems that prove the informational efficiency of the market.

KEY TERMS

invisible hand
competitive equilibrium model
double-auction market
Pareto efficient
first theorem of welfare economics
income inequality
deadweight loss

QUESTIONS FOR REVIEW

1. What are the information-processing, coordination, and motivation functions that arise when buyers and sellers interact?
2. Why is it difficult for one person or group of persons to perform the functions listed in question 1?
3. How does the market perform these functions?
4. How does the competitive equilibrium model explain the decisions of consumers and producers?
5. What is a double-auction market?
6. Do experimental double-auction markets validate the competitive equilibrium model?
7. What is the meaning of Pareto efficiency, and how does it differ from informational efficiency?

8. Why must marginal benefit equal marginal cost for Pareto efficiency?
9. Why is the sum of consumer surplus and producer surplus maximized in the market?
10. What is deadweight loss, and how do taxes cause it?

PROBLEMS

1. Suppose poor weather results in major damage to the coffee crop, resulting in lower supply.
   a. Draw a supply and demand diagram to show what will happen to the equilibrium price and quantity of coffee in the United States. Assume that the demand curve does not shift.
   b. Suppose the U.S. government observes that the price of coffee is increasing rapidly and imposes a price ceiling equal to the original equilibrium price. What effect does the price ceiling have on the quantity supplied and demanded of coffee? As a result of the price ceiling, how much coffee will actually be bought and sold?
   c. How are consumer and producer surplus affected by the price ceiling?
2. In 1975, 18 million calculators were produced and sold at an average price of $60. In 1983, 31 million calculators were produced and sold at an average price of $30. Assume that the demand curve for calculators did not shift between 1975 and 1983.
   a. Sketch market demand and market supply curves for calculators in 1975 and in 1983 to illustrate the change in price and quantity. Mark the 1975 and 1983 prices and quantities on the axes.
   b. Describe an event that could have led to the changes you just illustrated.
   c. Show the gain to consumers from this event on your sketch.
3. Consider the following supply and demand schedule:

Price	Supply	Demand
$.25	2	14
$.50	6	12
$.75	10	10
$1.00	14	8
$1.25	18	6
$1.50	22	4
$1.75	26	2

   a. Sketch the market supply and demand curves. Show the equilibrium quantity, price, producer surplus, and consumer surplus.
   b. Describe what would happen to the price of this product if a tax of $.75 per unit sold were enacted by the government. Show your answer graphically.
   c. Show the deadweight loss due to the tax on your diagram.
4. Calculate the consumer surplus using the market demand curve in Figure 7.3. Assume that the market price is $10. Show that you get the same answer by adding up the consumer surplus for all five buyers. How much does consumer surplus increase for the market as a whole and for each individual when the market price falls to $5?
5. Calculate the producer surplus using the market supply curve in Figure 7.4. Assume that the market price is $10. Show that you get the same answer by adding up the producer surplus for all five sellers. How much does producer surplus increase for the market as a whole and for each seller when the market price rises to $15?
6. Suppose that in a competitive market there are three buyers (Linda, Sue, and Pete) with the marginal benefit (*MB*) schedules below. If the price is $8, what will be the consumer surplus for each person? What is the consumer surplus for the market as a whole?

Quantity	*MB*—Linda	*MB*—Sue	*MB*—Pete
1	15	14	13
2	12	11	10
3	9	8	7
4	6	5	4
5	3	2	1

7. Suppose that in a competitive market there are three sellers (Max, Scott, and Karen) with the marginal cost (*MC*) schedules shown below. If the price is $8, what will be the producer surplus for each person? What is the producer surplus for the market as a whole?

Quantity	*MC*—Max	*MC*—Scott	*MC*—Karen
1	3	2	1
2	6	5	4
3	9	8	7
4	12	11	10
5	15	14	13

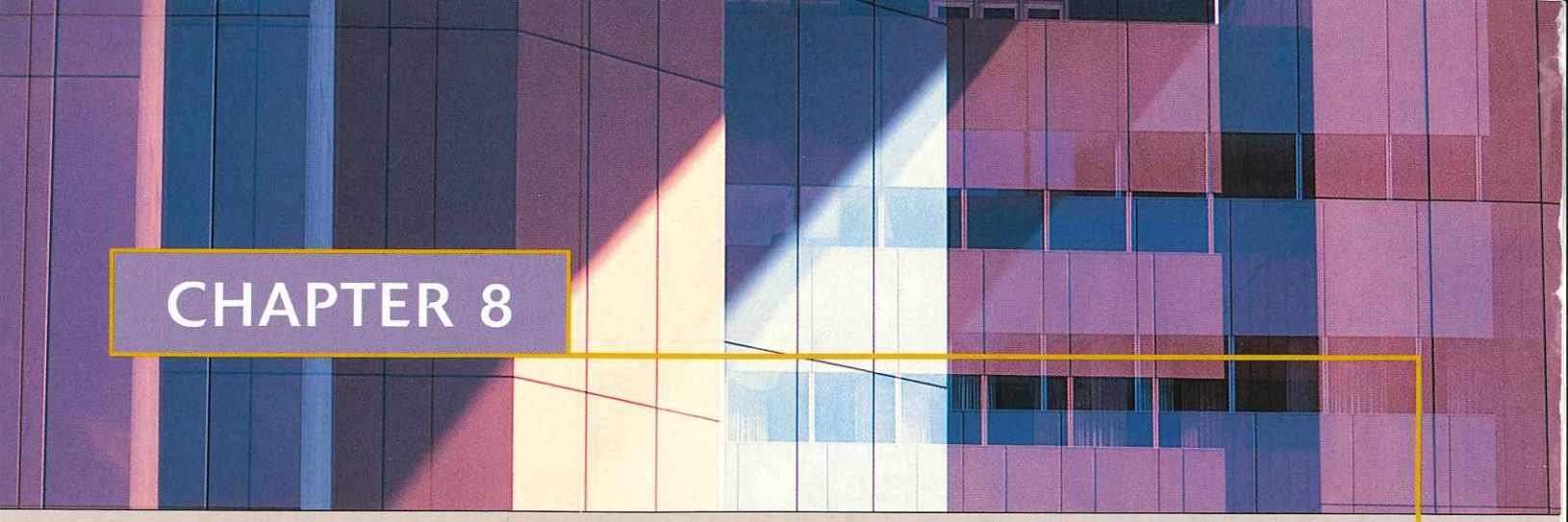
8. Suppose that in the market from the previous two questions the government imposes a $4 sales tax, which causes the equilibrium price to go up to $10. Draw a graph showing the original supply curve, the new supply curve, and the demand curve. Use this graph to show:
   a. The deadweight loss resulting from the $4 tax based on the original price and quantity.
   b. The amount of revenue collected by the government.
   c. The effect the tax has on producer and consumer surplus.
9. For which of the following items—milk, wine, coffee, bread, gasoline, 100 percent wool sweaters, sports cars, VCRs—do you think the deadweight loss of a sales tax would be the largest? The smallest? Explain.

10. What would the competitive equilibrium model predict about the quantity sold if people were not allowed to bid or ask more than $4 for any good in the market described in problems 4 and 5? What if they were not allowed to bid or ask less than $12? Illustrate your answers in a graph and show the deadweight loss in each case.

11. Firm A and firm B both produce the same product with the following total costs:

Firm A		Firm B	
Quantity Produced	*Total Costs*	*Quantity Produced*	*Total Costs*
0	5	0	2
1	6	1	5
2	8	2	9
3	11	3	14
4	15	4	20

Consider a situation in which 4 units are produced: Firm A produces 2 units, and firm B produces 2 units. Explain why this situation is not Pareto efficient. Could such a situation occur in a competitive market if both firms maximized profits? How could production be changed at the two firms in order to produce the 4 items at lower cost? Suppose the price is $3. How much would each firm produce?



CHAPTER 8

Costs and the Changes at Firms over Time

On a cold Saturday morning in December 1999, a wrenching story appeared on the front page of a California newspaper. It began:

> The end came at precisely 11:46 a.m. Friday. After 82 years, after four generations of toil and take-home pay, with deep roots tapping into two centuries, the end came without frill or fanfare. . . . There were tears. There were handshakes. . . . "Listen. There's a silence. It's like a hush has fallen over the place," said Bob Armstrong, the superintendent of the plant, which has been cranking out cans since the United States entered World War II in 1941.[1]

The story was an account of a firm—a cannery—shutting down its production facilities. Many firms, such as this cannery, shut down each year in the United States. But many more firms start up, so that the number of firms in existence continues to increase year after year. In the United States, about 700,000 firms

1. Geoffrey Tomb, "As the Final Harvest Ends, a Tech Torture Begins," *San Jose Mercury News*, December 18, 1999, p. 1.

Small businesses make up the largest share of businesses in the United States, employing, on average, half of all private sector employees. There are nearly 23 million small businesses in the United States, and they are opening and closing at a fairly steady rate. At least two-thirds of small businesses survive the first two years; about fifty percent are still open after four years. Small businesses close for a variety of reasons, including lack of adequate capital, but some businesses simply run their course. This storefront sign announces the 2002 closing of Florence Huie's laundry business in Derby, Connecticut—a business Ms. Huie started with her husband in 1960.

start up each year. Many firms are successful and will expand, but others may have to downsize and eventually shut down.

The purpose of this chapter is to develop a model for analyzing such changes at firms over time. To do so, we will extend the model of firm behavior developed in Chapter 6.

Costs, as we will see, are of vital importance to a firm's decision to start up, expand, or shut down. We will look carefully at firms' costs in this chapter. Some of the most rapidly growing firms in the United States have prospered because of new technologies that cut costs. For example, Wal-Mart—a rapidly expanding firm in the 1990s—developed a system whereby salesclerks electronically scan a bar code on each item purchased and automatically transmit the information back to the manufacturer, who can then immediately begin producing more of that item. This reduced costs.

Costs determine how large firms should be. Differences in the costs of manufacturing cement and haircuts, for example, mean that cement firms are usually large and barbershops are usually small. Costs also determine how firms should expand. For example, when firms choose between expanding their manufacturing facilities in the United States and acquiring another company abroad, they take into account the costs of labor and transportation as well as the effects of government policy toward firms.

Costs are such a crucial determinant of firm behavior that economists can capture the whole essence of a firm with a graph of its costs. By looking at such a graph, economists can determine the profitability of a firm and whether it should shut down or expand. This chapter shows how.

Finding Average Cost at an Individual Firm

In this section, we show how to find the average cost of production at a hypothetical transportation firm called On-the-Move. We will see that *average cost* plays an important role in the firm's decisions. To understand average cost, we must first review the ideas of total costs, fixed costs, variable costs, and marginal cost and show how these are related to the firm's production function.

Total Costs, Fixed Costs, Variable Costs, and Marginal Cost

total costs: the sum of variable costs and fixed costs. (Ch. 6)

fixed costs: costs of production that do not depend on the quantity of production. (Ch. 6)

variable costs: costs of production that vary with the quantity of production. (Ch. 6)

Total costs (*TC*) are the sum of all costs incurred by a firm in producing goods or services. The more that is produced, the larger are total costs. Recall from Chapter 6 that *fixed costs* (*FC*) and *variable costs* (*VC*) are the two key components of total costs.

Fixed costs are the part of total costs that do not vary with the amount produced in the short run; fixed costs include the cost of the factories, land, machines, and all other things that do not change when production changes in the short run. *Variable costs* are the part of total costs that vary in the short run as production changes. Variable costs include wage payments for workers, gasoline for trucks, fertilizer for crops, and all other things that change when the amount produced changes. By definition, total costs equal fixed costs plus variable costs; or, in symbols, $TC = FC + VC$.

short run: the period of time during which it is not possible to change all inputs to production; only some inputs, such as labor, can be changed.

long run: the minimum period of time during which all inputs to production can be changed.

■ **The Short Run and the Long Run.** Distinguishing the short run from the long run is the key to distinguishing fixed costs from variable costs. The *short run* and the *long run* are two broad categories into which economists parcel time. The **short run** is the period of time during which it is not possible to change all the inputs to production; only some inputs, such as labor, can be changed in the short run. The short run is too short, for example, to build a new factory or apartment building, to lay a fiber-optic cable, to launch a new communications satellite, or to get out of a lease on a storefront. The **long run,** in contrast, is long enough that all inputs, including capital, can be changed. Hence, the cost of each of the items that cannot be changed in the short run—factories, buildings, satellites—is fixed in the short run but can be changed in the long run.

Economists frequently use *capital* as an example of a factor that does not change in the short run and use *labor* as an example of a factor that can change in the short run. In fact, salaries paid to certain types of workers who have special skills and knowledge about the firm are better viewed as being fixed costs, and rents on certain capital items such as laptop computers or sewing machines are better viewed as being variable costs. Nevertheless, in the examples in this chapter, we refer to the cost of labor as the main variable cost and the cost of capital as the main fixed cost.

■ **Costs for On-the-Move.** Table 8.1 illustrates these definitions with cost data for On-the-Move. The firm, located in Houston, Texas, specializes in the strenuous but delicate job of moving pianos from one part of Houston to another. We use these hypothetical data rather than actual data to keep the example simple, but it is important to realize that the same analysis can be applied to data from any firm. Roadway Express, an actual moving firm that started in Houston with 16 trucks and has since gone nationwide, is a more complex example illustrating the same point. Table 8.1 lists the total costs, fixed costs, and variable costs for different levels of output at On-the-Move. Observe that fixed costs do not change but variable costs increase with output.

Table 8.1
Finding Average and Marginal Cost for On-the-Move
(costs measured in dollars per day)

Quantity (pianos moved per day) (*Q*)	Total Costs (*TC*)	Fixed Costs (*FC*)	Variable Costs (*VC*)	Average Total Cost (*ATC*)	Average Fixed Cost (*AFC*)	Average Variable Cost (*AVC*)	Marginal Cost (*MC*)
0	300	300	0	—	—	—	—
1	450	300	150	450	300	150	150
2	570	300	270	285	150	135	120
3	670	300	370	223	100	123	100
4	780	300	480	195	75	120	110
5	900	300	600	180	60	120	120
6	1,040	300	740	173	50	123	140
7	1,200	300	900	171	43	128	160
8	1,390	300	1,090	174	38	136	190
9	1,640	300	1,340	182	33	149	250
10	1,960	300	1,660	196	30	166	320
11	2,460	300	2,160	223	27	196	500

$TC = FC + VC$ $\quad ATC = \frac{TC}{Q}$ $\quad AFC = \frac{FC}{Q}$ $\quad AVC = \frac{VC}{Q}$ $\quad \frac{\text{Change in } TC}{\text{Change in } Q}$

The pictographs in Figure 8.1 show that fixed costs do not change in the short run at On-the-Move. Fixed costs, or the cost for four trucks and two terminals where the trucks are parked, are $300 per day regardless of how many pianos are moved during the day. Figure 8.1 also shows that variable costs increase with the amount produced. They increase from $600 to $1,660 as the number of pianos delivered per day rises from 5 to 10. Variable costs are shown in Figure 8.1 to rise because additional workers are hired to carry the goods and to drive and service the trucks. Thus, total costs rise from $900 to $1,960 as the number of pianos delivered rises from 5 to 10.

Figure 8.2 shows the same type of information as Figure 8.1 in graph form. Pairs of numbers on total costs and quantity from Table 8.1 are plotted in Figure 8.2. Connecting these dots results in the total costs curve. You can see how the total costs of moving the pianos steadily increase with the number of pianos moved. Fixed costs are shown to be unchanged at all levels of output. Figure 8.2 shows variable costs by the distance between the total costs curve and the fixed costs curve.

marginal cost: the change in total costs due to a one-unit change in quantity produced. (Ch. 6)

■ **Marginal Cost.** Table 8.1 also shows how the *marginal cost* of On-the-Move depends on the quantity of services produced (the number of pianos moved). Recall from Chapter 6 that marginal cost is the change in total costs due to a one-unit change in the quantity produced. For example, the marginal cost of increasing production from 5 moves to 6 moves a day is $140, or the change in total costs ($1,040 − $900 = $140) divided by the change in production (6 − 5 = 1). The last column of Table 8.1 shows the marginal cost for each additional piano moved by On-the-Move, from the first to the eleventh piano.

Observe that in Table 8.1, marginal cost declines at low levels of production and then begins to increase again. Marginal cost reaches a minimum of $100 when production increases from 2 to 3 units of output. Recall that in the examples in Chapter 6

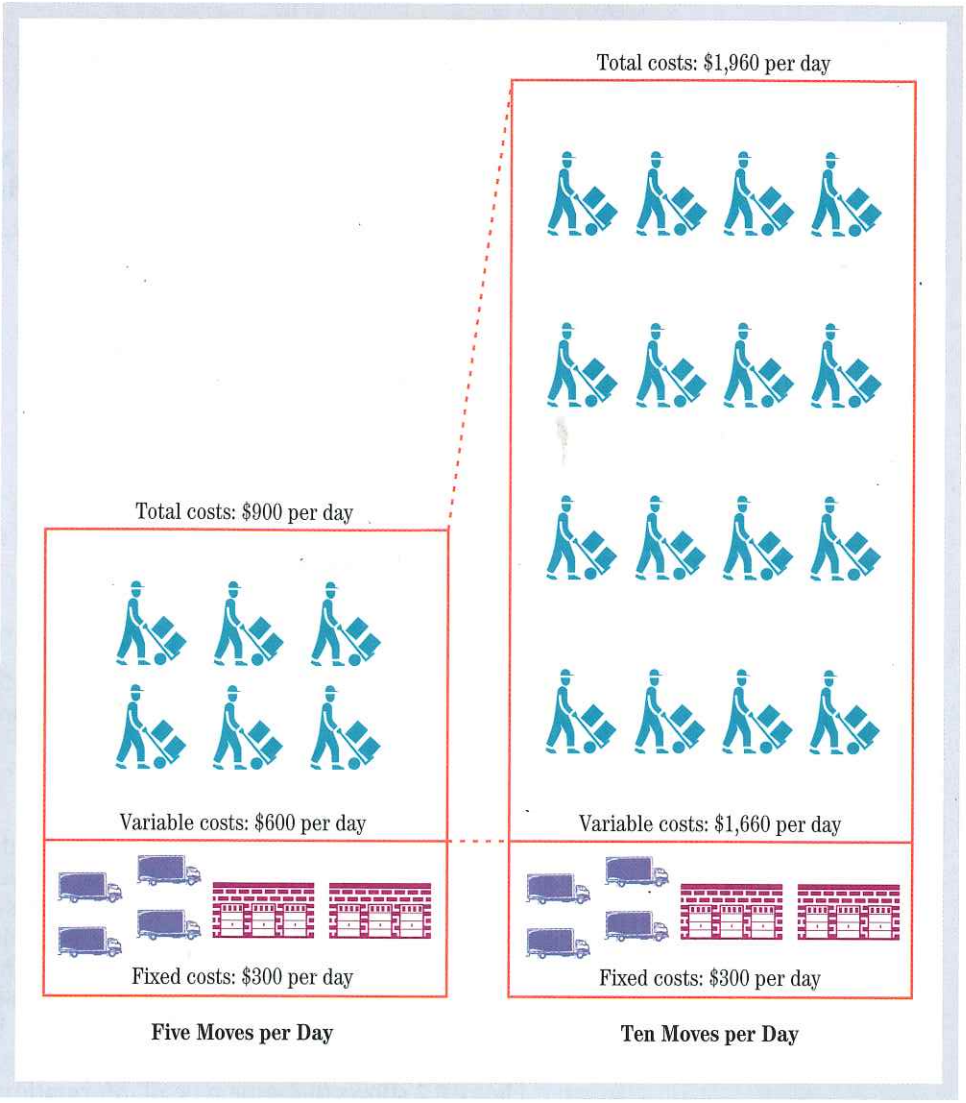


Figure 8.1
Fixed Costs versus Variable Costs
Fixed costs remain constant as the output of the firm increases in the short run. In the example of On-the-Move, fixed costs are the daily rental or interest costs for trucks and terminals under long-term lease or owned by the firm. Variable costs change with the level of output. In the case of On-the-Move, more workers must be hired to move more pianos.

(Table 6.4), marginal cost increased throughout the whole range of production. In the example of On-the-Move, marginal cost declines over part of the range of production. We will explain the reason for the difference, but first we need to define average cost.

average total cost (*ATC*): total costs of production divided by the quantity produced (also called cost per unit).

average variable cost (*AVC*): variable costs divided by the quantity produced.

average fixed cost (*AFC*): fixed costs divided by the quantity produced.

Average Cost

Average total cost (*ATC*) is defined as total costs (TC) of production divided by the quantity (Q) produced. In symbols, $ATC = TC/Q$. For example, if the total costs of producing 4 items are \$3,000, then the average total cost is \$750 (\$3,000/4). Another name for average total cost is *cost per unit*. We can also define average cost for fixed and variable costs. Thus, **average variable cost (*AVC*)** is defined as variable costs divided by the quantity produced: $AVC = VC/Q$. **Average fixed cost (*AFC*)** is defined as fixed costs divided by the quantity produced: $AFC = FC/Q$. Of the three averages,

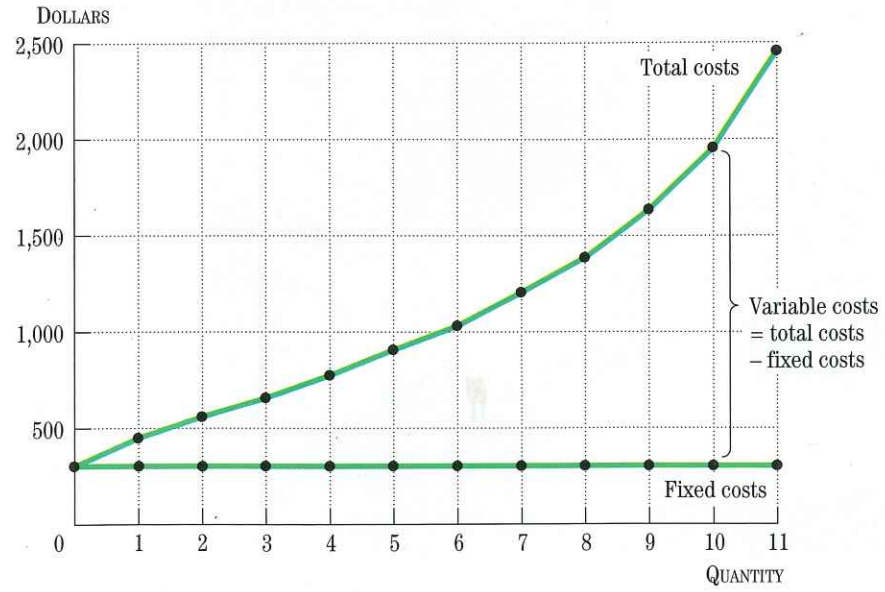


Figure 8.2
Total Costs Minus Fixed Costs Equal Variable Costs
The two lines on the diagram show total costs and fixed costs for On-the-Move. Variable costs are the difference between the two lines. Variable costs rise with production, but fixed costs are constant.

we will use average total cost most frequently; the other two averages are important for knowing whether to shut down a firm or keep it open when it is losing money.

Average total cost for On-the-Move is shown in Table 8.1. For example, total costs for 2 pianos moved ($Q = 2$) are \$570; dividing \$570 by 2 gives an average total cost of \$285. For 3 pianos moved ($Q = 3$), total costs are \$670; dividing by 3 gives \$223 for average total cost. Notice that average total cost declines as production increases from low levels. Then average total cost starts to increase. In the example, average total cost starts to increase at 8 units: \$1,390 divided by 8 is \$174, and \$1,200 divided by 7 is \$171. That average total cost first decreases and then increases as production rises is a common pattern for most firms.

Average variable cost is also illustrated in Table 8.1. For 2 pianos moved ($Q = 2$), for example, average variable cost is \$270 divided by 2, or \$135. You can see that average variable cost, in this example, first declines and then increases throughout the rest of the range of production.

Finally, observe in Table 8.1 that average fixed cost gets smaller as production rises. Because average fixed cost is calculated by dividing fixed costs by the quantity produced, average fixed cost must decline as the quantity produced rises.

Costs Depend on the Firm's Production Function

The cost information in Table 8.1 is determined by how much *input* of labor and capital it takes to produce *output* and by the price of capital and labor. First consider some illustrative calculations of costs as the firm increases production from Q equals zero to 1 and then to 2.

■ **Varying Labor Input but Not Capital Input in the Short Run.** According to Table 8.1, it costs On-the-Move \$300 a day for capital, which is 4 trucks and 2 terminals. Let's assume that the \$300 consists of \$25 per day for each of the 4 trucks and

Inputs and the Production Function
Labor (the two workers) and capital (the truck) are inputs to production (moving the piano).

$100 per day for each of the 2 terminals ($25 × 4 + $100 × 2 = $300). These are the fixed costs that will be incurred even if zero pianos are moved. If the trucks and terminals were leased for one year, then the fixed costs would include the rental payment on the lease. If the trucks and the terminals were purchased on credit by On-the-Move, then the fixed costs would include interest payments on the loans. If the trucks and the terminals were bought outright, then the fixed costs would include the opportunity cost—the forgone interest payments—of the funds used to buy the trucks and the terminals.

To move pianos, however, On-the-Move needs labor. To move 1 piano, it might be enough to have 1 driver, 1 mechanic to service the truck, and another worker to help carry and load the piano. The example assumes that the cost of labor input is $150, which might consist of 15 hours of work at $10 per hour; perhaps 5 hours of work for each of the 3 workers. As production increases from $Q = 0$ to $Q = 1$, variable costs increase from zero to $150 and total costs increase from $300 to $450. Thus, the marginal cost of moving 1 piano rather than zero pianos is $150.

To move to a higher level of production, On-the-Move requires more workers. According to Table 8.1, if production rises from 1 piano moved to 2 pianos moved, then total costs increase from $450 to $570; marginal cost is $120. With wages of $10 an hour, this marginal cost would be the cost of 12 more hours of work—perhaps another driver and loader each working 5 hours a day plus increasing the hours of the mechanic by 2. As we observed already, marginal cost *declines* as production increases from 1 to 2 units of output. Now we are beginning to see why. Marginal cost decreases because labor input rises by less when increasing production from 1 to 2 units than it does when increasing production from zero to 1 unit. The reason has to do with the nature of the firm's production; perhaps the mechanic can service two trucks in less than twice the time it takes to service one truck (5 hours for one, 7 hours for two). Although these calculations illustrate how costs depend on the inputs to production, to see what is going on throughout the whole range of production, we need to look at the firm's production function.

production function: a relationship that shows the quantity of output for any given amount of input. (Ch. 6)

■ **The Production Function.** Table 8.2 shows the number of hours of work required to move different numbers of pianos at On-the-Move. It is On-the-Move's short-run *production function*, showing how much output can be produced for each amount of labor input. To calculate the variable costs at On-the-Move, using the information in Table 8.2, continue to assume that the wage is $10 per hour. Then to move 1 piano takes 15 hours of work; at $10 per hour, variable costs are $150. To move 10 pianos takes 166 hours of work; at $10 per hour, variable costs are $1,660. Similar calculations for all levels of output are shown in the third column of Table 8.2. Note that the variable costs in Table 8.2 are the same as those in Table 8.1. Thus, we have shown explicitly how the firm's costs depend on its production function.

Recall from Chapter 6 that the *marginal product of labor* is the change in production that can be obtained with an additional unit of labor. Decreasing marginal product of labor is called *diminishing returns to labor*. Increasing marginal product of labor is called *increasing returns to labor*.

Table 8.2
Using the Production Function to Compute Variable Costs

Quantity (pianos moved)	Hours of Work	Labor Costs at $10 Wage (variable costs)
0	0	0
1	15	150
2	27	270
3	37	370
4	48	480
5	60	600
6	74	740
7	90	900
8	109	1,090
9	134	1,340
10	166	1,660
11	216	2,160

Observe that increasing marginal product of labor exists at low levels of production; for example, it takes only 10 hours of labor to increase production [by 1 unit] from 2 to 3 units, whereas it takes 12 hours of labor to increase production [by 1 unit] from 1 to 2 units. At higher levels of production, decreasing marginal product of labor exists.

The marginal product of labor is illustrated in Figure 8.3, which shows a graph of the production function from Table 8.2. Figure 8.3 shows how, for low levels of labor input, the marginal product of labor increases: Output increases by more from a given change in labor input as labor input increases. At low levels of production, increasing marginal product of labor is a possibility because the firm's capital can be better utilized: During an oil change, for example, the mechanic can work on another truck as oil drains from the first truck. At high levels of labor input, marginal product starts to decline; diminishing returns set in. The same increase in labor input results in smaller and smaller increases in output.

Table 8.2 shows increasing marginal product of labor up to 3 units of output produced. Then diminishing returns begin. This pattern of increasing marginal product of labor up to 3 units of output and then decreasing marginal product is what causes the pattern of decreasing marginal cost up to 3 units of output produced followed by increasing marginal cost. More generally, if the marginal product of labor is increasing, then marginal cost is decreasing, and vice versa. With more and more workers required to produce a given amount of output, the marginal cost of producing that amount will increase. To summarize:

Increasing marginal product of labor → Decreasing marginal cost

Decreasing marginal product of labor → Increasing marginal cost

average product of labor: the quantity produced divided by the amount of labor input.

Be sure to distinguish between the marginal product of labor and the average product of labor. The **average product of labor** is the quantity produced, or *total product*, divided by the amount of labor input. Thus, the average product of labor is Q/L, where Q is total product and L is labor input. On the other hand, the marginal product of labor is $\Delta Q/\Delta L$, where ΔQ is the change in the quantity produced and ΔL is the change in labor input.

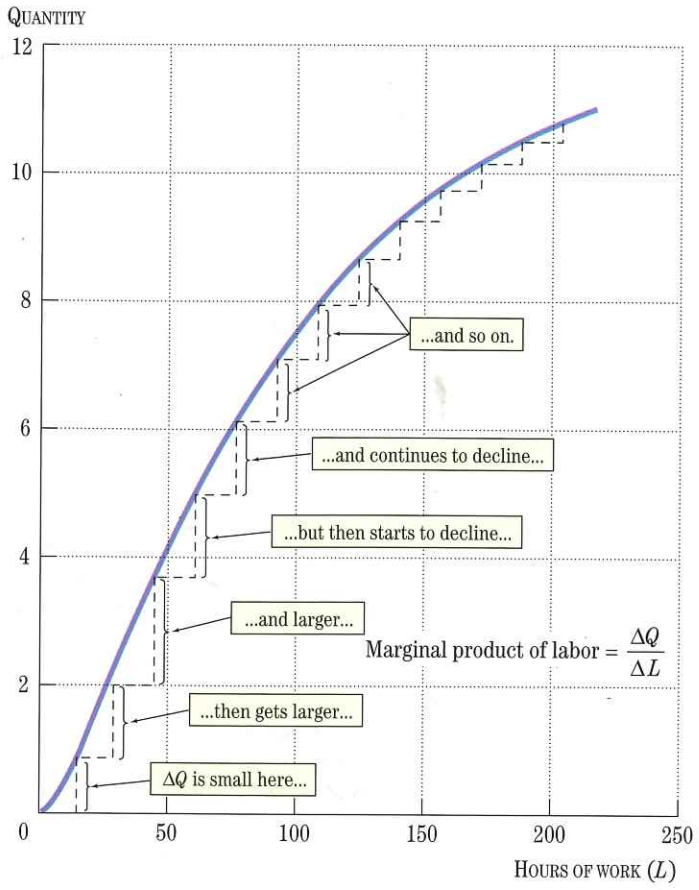


Figure 8.3
On-the-Move's Production Function
The curve shows the production function in which more labor input gives more output. Capital (trucks and terminals) is not changed. Observe that the marginal product of labor first increases and then decreases with more labor input.

REVIEW

- The short run is the period of time in which it is not possible to change all inputs to production; only some inputs, such as labor, can be changed. The long run is the period of time in which the firm can vary all inputs to production, including capital.
- Fixed costs are constant in the short run for all levels of production. Variable costs increase in the short run as more is produced. Total costs are fixed costs plus variable costs.
- Useful information about a firm comes from looking at average cost, or cost per unit. There are three types of average cost:
  1. Average total cost is defined as total costs divided by quantity, or $ATC = TC/Q$.
  2. Average variable cost is defined as variable costs divided by quantity, or $AVC = VC/Q$.
  3. Average fixed cost is defined as fixed costs divided by quantity, or $AFC = FC/Q$.
- Costs depend on the firm's production function. When marginal product of labor is decreasing, marginal cost is increasing.

Average Cost Curves

The information about average cost in Table 8.1 can be turned into an informative graph, as shown in Figure 8.4. The vertical axis of Figure 8.4 shows the dollar cost, and the horizontal axis shows the quantity produced. The pairs of points from Table 8.1 are plotted as dots in Figure 8.4, and the dots have been connected to help visualize the curves. Although the curves use exactly the same information as in Table 8.1, they are more useful. The curves are called the *marginal cost curve*, the *average total cost curve*, and the *average variable cost curve*. We label the curves *MC*, *ATC*, and *AVC*, respectively.

It is very clear from Figure 8.4 that marginal cost first decreases and then increases, as observed in Table 8.1. We now know the reason: The marginal product of each additional worker increases at lower levels of production and then decreases at higher levels of production, as shown in Figure 8.3.

Figure 8.4 also makes it very clear that average total cost first declines and then increases. In other words, the average total cost curve is *U-shaped*.

The relative positions of the average total cost curve and the marginal cost curve in Figure 8.4 are important and will come up repeatedly. Observe that when the marginal cost curve is below the average total cost curve, average total cost is declining. For example, when production rises from $Q = 1$ to $Q = 2$ in Figure 8.4, marginal cost is less than average total cost and average total cost declines. However, look at the right-hand side of Figure 8.4, where marginal cost is greater than average total cost; then average total cost increases. For example, a marginal cost of \$320 is greater than the average total cost of \$182, and average total cost goes up to \$196. This is a general and important result: *When marginal cost is less than average total cost, then average total cost is declining; when marginal cost is greater than average total cost, then average total cost is increasing.*

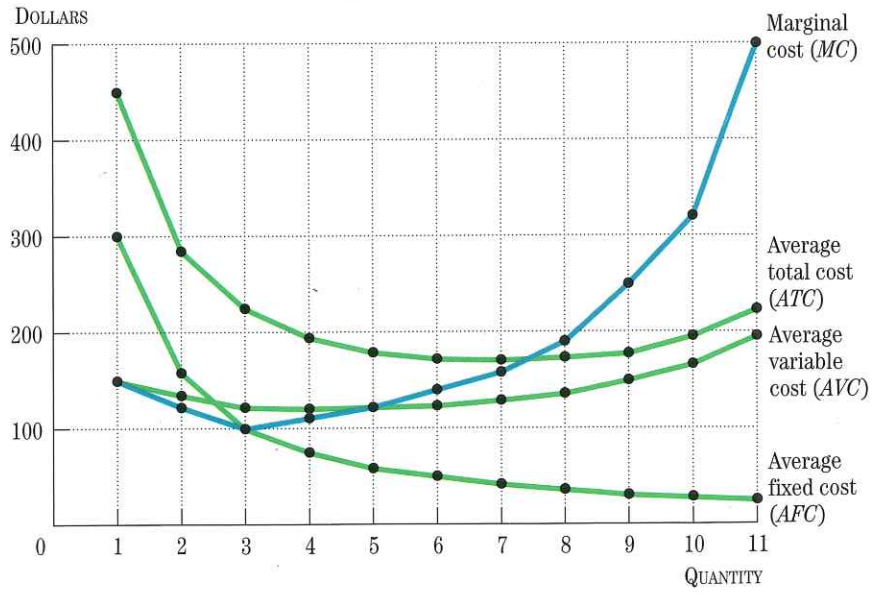


Figure 8.4
Average Cost and Marginal Cost from a Numerical Example
Average total cost first declines then increases as more is produced. Marginal cost is below average total cost when average total cost is falling and above average total cost when average total cost is rising. This relationship also holds between average variable cost and marginal cost. These cost curves are plotted from the data given in Table 8.1.

This result also holds for average variable cost: If marginal cost is greater than average variable cost, then average variable cost is increasing; if marginal cost is less than average variable cost, then average variable cost is decreasing. These relationships between the two average cost curves and the marginal cost curve are essential to the analysis that follows.

Marginal versus Average in the Classroom

The reason for the relationship between marginal cost and average total cost or average variable cost can be seen with an analogy. Consider another example of averages, say, average grades on the midterm exam in an economics class. Suppose that the average grade of people in the classroom the day after an exam is 64. Now imagine that another person, who has a midterm grade of 100, enters the classroom. We know that 100 is greater than the average grade of 64. In other words, the "marginal grade" of 100 is greater than the average grade; when the person with the grade of 100 enters the room, the average grade in the classroom increases. Now suppose that a different person who has never attended any lectures and has a midterm grade of zero comes in. The "marginal grade" of zero is less than the average grade of 64; hence, the average grade declines. This is a property of averaging and applies to grades, heights, weights, and so on, as well as to costs. When you bring someone into a group whose grade is less than the group's average, then the average declines. A below-average contribution causes the average to fall; on the other hand, an above-average contribution causes the average to rise. The relationships between marginal cost and average total cost or average variable cost say nothing more than this.

Generic Cost Curves

The relationship between marginal and average allows us to sketch a *generic* cost curve diagram, the general properties of which characterize virtually all firms, not just On-the-Move. Such a diagram is shown in Figure 8.5. Again, the vertical axis is the dollar cost and the horizontal axis is the quantity, but in a generic picture, we do not scale the axes because they apply to any firm, whether in textiles, moving, or electronics. Note that the marginal cost curve cuts both the average variable cost curve and the average total cost curve at their minimum points. To the left of the point where the curves cross, marginal cost is less than average total cost, and so average total cost declines. As the marginal cost curve passes through the minimum, average total cost begins to increase. Try drawing your own diagram. If the marginal cost curve does not go through the lowest point of both the average total cost curve and the average variable cost curve, you have made an error.

There is another important relationship in Figure 8.5. The distance between the average total cost curve and the average variable cost curve gets smaller as production increases because fixed costs are a smaller and smaller proportion of total costs as production increases. Recall that fixed costs are the difference between total costs and variable costs. Thus, the gap between average total cost and average variable cost is average fixed cost, or fixed costs divided by quantity, FC/Q. Since fixed costs (FC) do not change, the ratio FC/Q declines as Q increases. The distance between the average total cost curve and the average variable cost curve is this distance FC/Q, which declines as quantity increases. Hence, the distance between the *ATC* curve and the *AVC* curve grows smaller as you move to the right in the diagram. Any picture you draw should show this relationship.

Observe that the marginal cost curve in the generic picture of Figure 8.5 has a region of declining marginal cost at low production levels. The graph allows for the

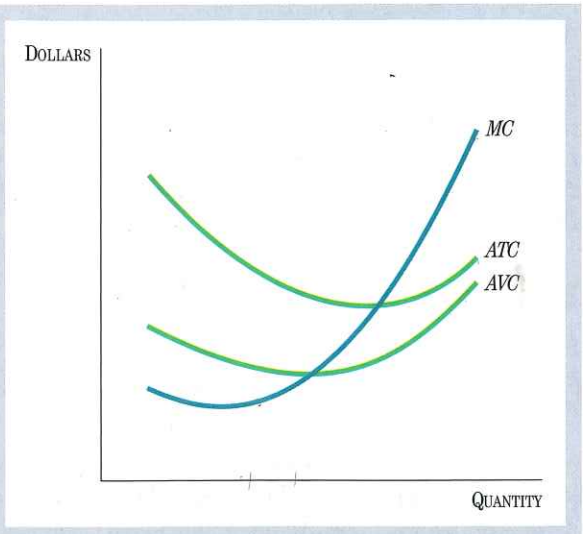


Figure 8.5
Generic Sketch of Average Cost and Marginal Cost
Every firm can be described by cost curves of the type drawn here. Compare these generic curves with the specific curves in Figure 8.4. Check these curves against the checklist in the margin.

It's helpful to use the following checklist when you draw this graph:

1. Make sure the marginal cost curve cuts through the average total cost curve and the average variable cost curve at their minimum points, and understand the reason for this.
2. Make sure the distance between average total cost and average variable cost gets smaller as you increase the amount of production.
3. Put a small dip on the left-hand side of the marginal cost curve before the upward slope begins. This makes your curve look more interesting and allows for the possibility of decreasing marginal cost at very low levels of production.

possibility that at low production levels, the marginal product of labor increases and, therefore, marginal cost declines. This was true for On-the-Move, which had increasing marginal product of labor up to 3 pianos moved, and we allow for it in the generic case.

You may have noticed that for the cost curves for On-the-Move in Figure 8.4, the marginal cost curve and the average variable cost curve touch at 1 unit of output. This occurs because the marginal cost of producing 1 rather than zero units of output must equal the variable cost of producing 1 unit, as shown in Table 8.1. Thus, if the generic cost curve were drawn all the way over to 1 unit of output on the left of Figure 8.5, the marginal cost and the average variable cost curve would start at the same point. Because we do not usually draw generic cost curves that go all the way over to the vertical axis, we do not usually show them starting at the same point.

REVIEW

- The marginal cost curve and the average cost curves are closely related. The marginal cost (*MC*) curve cuts through both the average total cost (*ATC*) curve and the average variable cost (*AVC*) curve at their lowest points.
- Another important property of a cost curve diagram is that the gap between average total cost and average variable cost gets smaller as more is produced.
- The relationships between marginal cost and the two average cost curves are represented by the following general rule: When marginal cost is less than average total cost (or average variable cost), then average total cost (or average variable cost) is declining; when marginal cost is greater than average total cost (or average variable cost), then average total cost (or average variable cost) is increasing.

Costs and Production: The Short Run

As we saw in Chapter 6, a competitive firm takes the market price as given. If it is maximizing profits, it will choose a quantity to produce in the short run such that its marginal cost equals the market price ($P = MC$). The resulting level of production for a competitive firm with the cost curves in Figure 8.5 is shown in Figure 8.6. The quantity produced is determined by the intersection of the marginal cost (MC) curve and the market price line (P). We draw a dashed vertical line to mark the quantity (Q) produced. But when the firm produces this quantity, are its profits positive, or is the firm running a loss? If it is running a loss, should it shut down in the short run? To answer these questions, we need to use the cost curves to find the firm's profits.

The Profit or Loss Rectangle

Profits equal total revenue minus total costs. To calculate profits with the average cost diagram, we need to represent total revenue and total costs on the average cost diagram.

■ **The Total Revenue Area.** Figure 8.6 shows a particular market price P and the corresponding level of production Q chosen by the firm. The total revenue that the firm gets from selling quantity Q is price P times quantity Q. Figure 8.7 shows that this total revenue can be represented by the area of a rectangle with width Q and height P. This rectangle is shown by the shaded area in Figure 8.7. Because the width

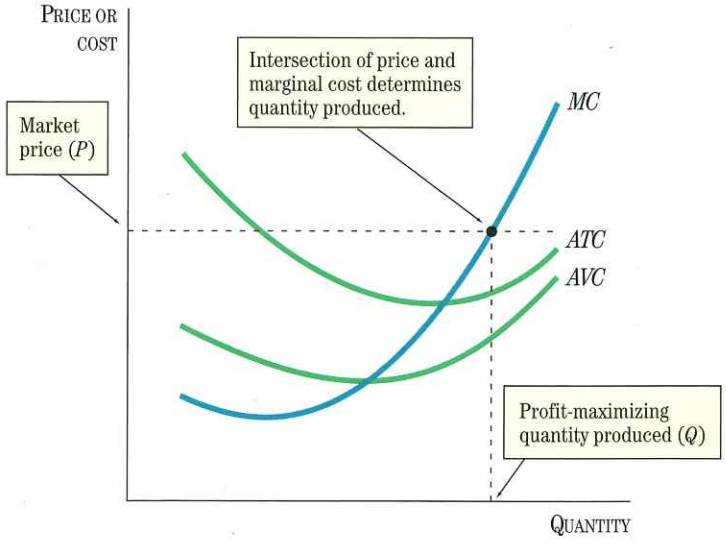


Figure 8.6
Price Equals Marginal Cost
If a firm is maximizing profits, then it chooses a quantity (Q) such that price equals marginal cost. Thus, the quantity is determined by the intersection of the market price line and the marginal cost curve, as shown on the diagram. In this picture the *ATC* and *AVC* curves are a sideshow, but they enter the main act in Figure 8.7, when we look at the firm's level of profits.

of this rectangle is the quantity produced Q and the height of this rectangle is the market price P, the area is $P \times Q$, or total revenue.

■ **The Total Costs Area.** Total costs can also be represented in Figure 8.7. First, observe the dashed vertical line in Figure 8.7 marking the profit-maximizing quan-tity produced. Next, observe the point where the average total cost curve intersects this dashed vertical line. This point tells us what the firm's average total cost is when it produces the profit-maximizing quantity. The area of the rectangle with the hash marks shows the firm's total costs. Why? Remember that average total cost is defined as total costs divided by quantity. If we take average total cost and multiply by quantity, we get total costs: $ATC \times Q = TC$. The quantity produced (Q) is the width of the rectangle, and average total cost (ATC) is the height of the rectangle. Hence, total costs are given by the area of the rectangle with the hash marks.

■ **Profits or Losses.** Since profits are total revenue less total costs, we compute profits by looking at the difference between the two rectangles. The difference is itself a rectangle, shown by the part of the revenue rectangle that rises above the total costs rectangle. *Profits are positive* because total revenue is greater than total costs in Figure 8.7. But profits can also be negative, as shown in Figure 8.8.

Suppose that the market price is at a point where the intersection of the marginal cost curve and the market price line gives a quantity of production for which average total cost is *above* the price. This situation is shown in Figure 8.8. At this lower price, we still have the necessary condition for profit maximization. The firm will produce the quantity that equates price and marginal cost, as shown by the intersection of the price line and the marginal cost curve.

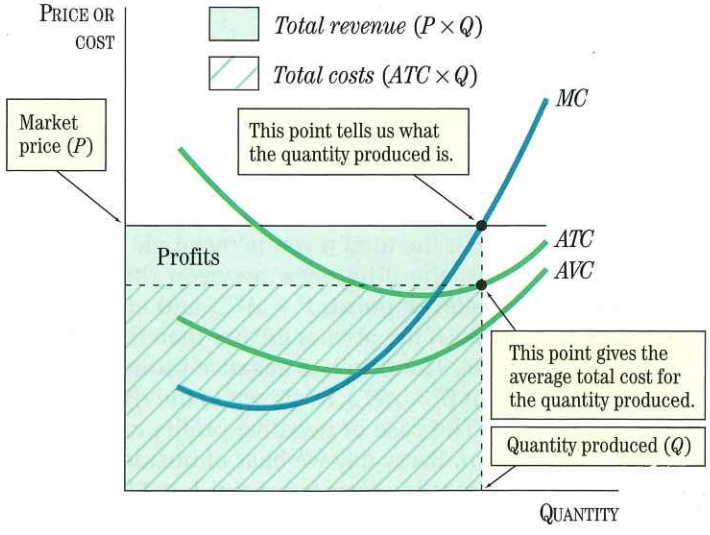


Figure 8.7
Showing Profits on the Cost Curve Diagram
The price and quantity produced are the same as those in Figure 8.6. The area of the shaded rectangle is total revenue. We use the *ATC* curve to find total costs in order to compute profits. First we mark where the *ATC* curve intersects the dashed vertical line showing the quantity produced. The area of the rectangle with the hash marks is total costs because the total costs (*TC*) equal average total cost (*ATC*) times quantity produced, $TC = ATC \times Q$. The part of the shaded rectangle rising above the hash-marked area is profits.

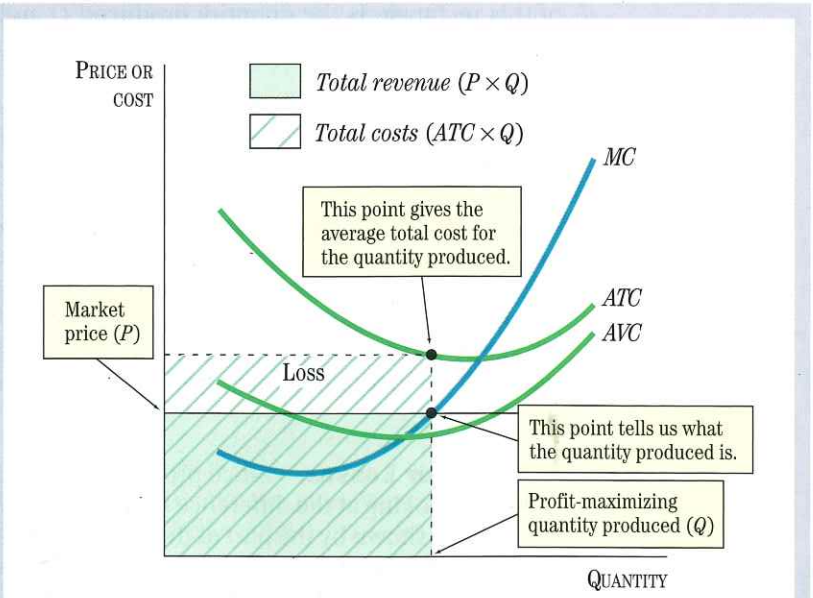


Figure 8.8
Showing a Loss on the Cost Curve Diagram
Here the market price is lower than in Figure 8.7. The market price line intersects the marginal cost curve at a point below the average total cost curve. Thus, the area of the total costs rectangle is larger than the area of the total revenue rectangle. Profits are less than zero, and the loss is shown in the diagram.

The amount of total revenue at this price is again price times quantity ($P \times Q$), or the shaded rectangle.

Total costs are average total cost times the quantity produced, that is, $ATC \times Q$, or the area of the rectangle with the hash marks.

The difference between total revenue and total costs is profit, but in this case *profits are negative*, or there is a loss. Total revenue is less than total costs, as shown by the cost rectangle's extending above the revenue rectangle. The extent of cost overhang is the loss.

The Breakeven Point

breakeven point: the point at which price equals the minimum of average total cost.

Now draw the market price line through the point where the marginal cost curve intersects the average total cost curve. Recall that this is the minimum point on the average total cost curve. This situation is shown in the middle panel of Figure 8.9. At that price, the firm chooses a quantity for which average total cost equals the price, so that the total revenue rectangle and the total cost rectangle are exactly the same. Thus, the difference between their areas is zero. At that price, the firm is at a **breakeven point:** $P = ATC$, and economic profits are zero. The firm earns positive profits if the price is greater than the breakeven point ($P > ATC$), as shown in the left panel. The firm has negative profits (a loss) if the price is lower than the breakeven point ($P < ATC$), as shown in the right panel of Figure 8.9.

The case of negative profits raises the question of why the firm does not shut down. Every day we hear of businesses losing money. In 1993, for example, Adidas, the running shoe company, lost $100 million, but it did not shut down. Why does a firm with negative profits stay in business? The reason is that if the firm shut down, losses would be even larger. In the short run, the fixed costs have to be paid. By continuing operations, the firm can minimize its losses. Let's examine this more carefully and determine when exactly the firm should shut down.

You can use some algebra to check the result that a firm should stop producing when the price is less than average variable cost: $P < AVC$. Note that

Profits = total revenue − total costs

Because total costs equal variable costs plus fixed costs, we can replace total costs to get

$$\text{Profits} = P \times Q - (VC + FC)$$

Now, since

$VC = AVC \times Q$, we have

$$\text{Profits} = P \times Q - AVC \times Q - FC$$

Rearranging this gives

$$\text{Profits} = (P - AVC) \times Q - FC$$

If $P < AVC$, the first term in this expression is negative unless $Q = 0$. Thus, if $P < AVC$, the best your firm can do is set $Q = 0$. This eliminates the negative drain on profits in the first term in the last expression. You minimize your loss by setting $Q = 0$.

The Shutdown Point

The firm should shut down if the price falls below the minimum point of the average variable cost curve and is not expected to rise again. In this case, the market price equals marginal cost at a quantity where total revenue ($P \times Q$) is smaller than the variable costs ($AVC \times Q$) of producing at that point.

When total revenue is less than variable costs, it makes sense to stop producing. For example, if the price of moving pianos is so low that the revenue from moving the pianos is less than the workers are paid to move the pianos, it is best not to move any pianos. It is better to shut down production. The fixed costs for the trucks and the garage would have been paid, but with the price so low, revenues cannot cover the payment to the workers.

The right panel of Figure 8.10 shows the case where the price is below average variable cost ($P < AVC$) and the firm should shut down. However, if the price is above average variable cost ($P > AVC$), as shown in the left panel of Figure 8.10, the firm should not shut down, even if the price is below average total cost and profits are negative. Because total revenue is greater than variable costs, shutting down would eliminate this extra revenue. It is better to keep producing in the short run. We assume that the firm must pay the fixed costs, as it is obligated for them over the short run. For example, Adidas did not shut down in 1993 because it had to pay fixed costs in the short run; with the price of running shoes greater than the average variable cost of producing them, the losses were less than if it had shut down.

Economists have developed the concept of sunk cost, which may help you understand and remember why a firm like Adidas would continue to operate in the short run even though it was reporting losses. A *sunk cost* is a cost that you have committed to pay and that you cannot recover. For example, if a firm signs a year's lease for factory space, it must make rental payments until the lease is up, whether the

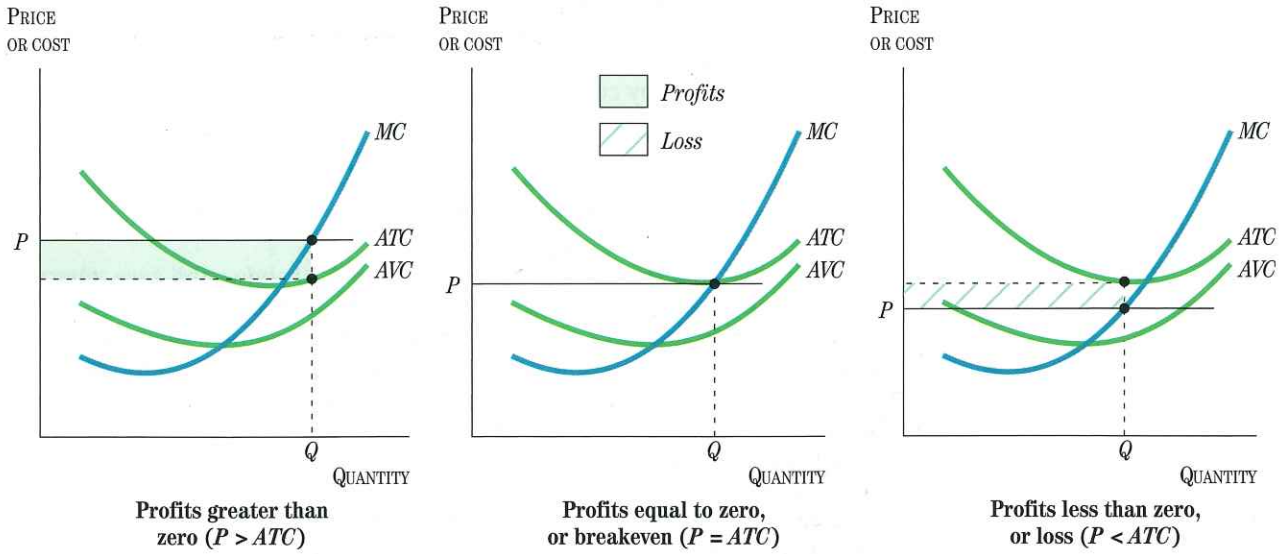


Figure 8.9
The Breakeven Point

When profits are zero, we are at the breakeven point, as shown in the middle panel. In this case, the market price line intersects the marginal cost curve exactly where it crosses the average total cost curve. The left panel shows a higher market price, and profits are greater than zero. The right panel shows a lower market price, and profits are less than zero—there is a loss. The cost curves are exactly the same in each diagram.

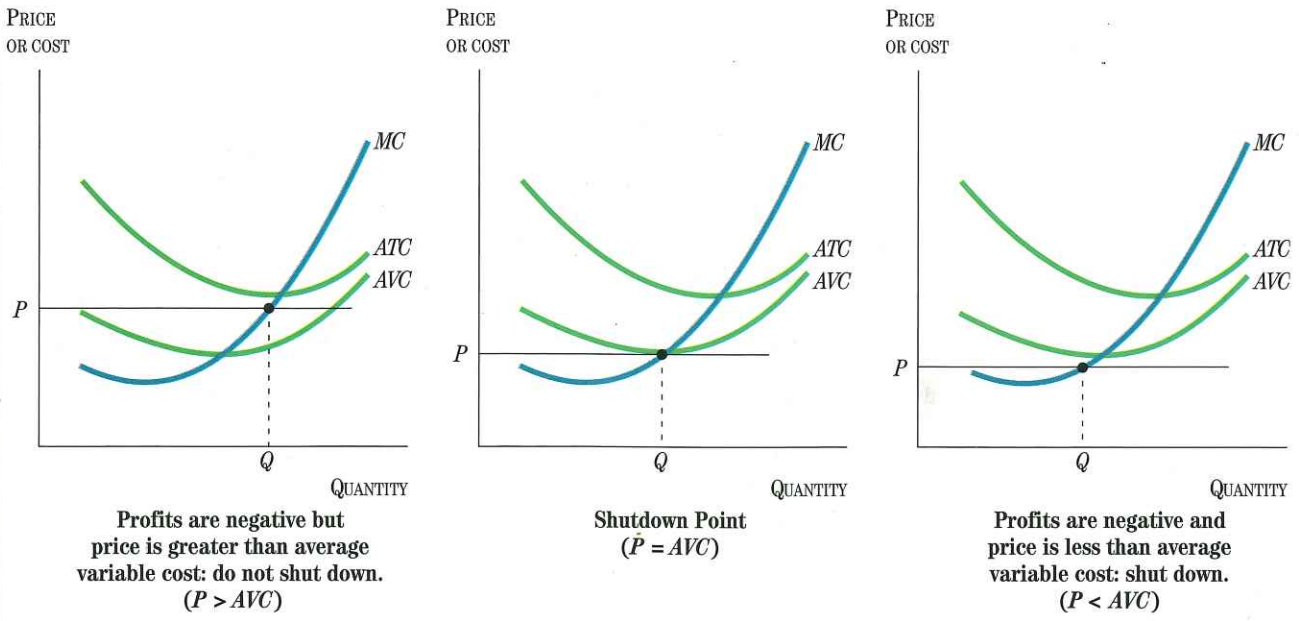


Figure 8.10
The Shutdown Point

When price is above average variable cost, the firm should keep producing even if profits are negative. But if price is less than average variable cost, losses will be smaller if the firm shuts down and stops producing. Hence, the shutdown point is when price equals average variable cost.

space is used or not. The rental payments are an example of a sunk cost. In the short run, fixed costs—such as lease payments—are sunk costs to a firm. The firm cannot recover these costs by shutting down. All Adidas could do in the short run was reduce its losses, and it did so by continuing to produce (assuming that the revenue from each pair of shoes was greater than the variable cost of producing them). The important thing about a sunk cost is that once you commit to it, there is nothing you can do about it, so you might as well ignore it in your decisions. If you paid for season tickets to the opera and then lost your taste for opera, there is no reason to go to the opera and be miserable.

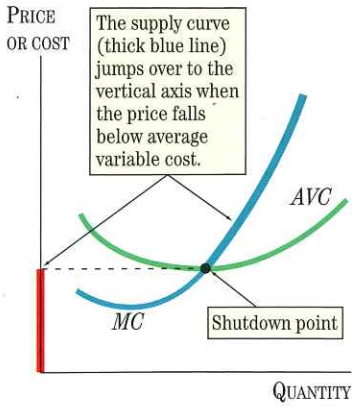
Now, observe that the middle panel of Figure 8.10 shows the case where price exactly equals the minimum point of the average variable cost curve. In this case, price equals marginal cost at a quantity where total revenue equals variable costs and price equals average variable cost ($P = AVC$). This is called the **shutdown point.** If the price falls below the shutdown point, the firm should shut down. If the price is above the shutdown point, the firm should continue producing.

shutdown point: the point at which price equals the minimum of average variable cost.

In thinking about the shutdown point, the time period is important. We are looking at the firm during the short run, when it is obligated to pay its fixed costs and cannot alter its capital. The question for On-the-Move is what to do when it has already committed to paying for the trucks and terminals, but the price of moving pianos falls to such a low level that it does not cover variable costs. The shutdown rule says to stop in that situation. However, if profits are negative and the price is greater than average variable cost, then it is best to keep producing.

Two different points:
Shutdown point
P = minimum AVC
Breakeven point
P = minimum ATC

The shutdown point can be incorporated into the firm's supply curve. Recall that the supply curve of a single firm tells us the quantity of a good that the firm will



produce at different prices. As long as the price is above average variable cost, the firm will produce a quantity such that marginal cost equals the price. Thus, for prices above average variable cost, the marginal cost curve is the firm's supply curve, as shown in the figure in the margin. However, if the price falls below average variable cost, then the firm will shut down; in other words, the quantity produced will equal zero ($Q = 0$). Thus, for prices below average variable cost, the supply curve jumps over to the vertical axis, where $Q = 0$, as shown in the figure in the margin.

Be sure that you understand the difference between these two points: the shutdown point and the breakeven point. The shutdown point is where price equals the minimum average variable cost (Figure 8.10). The breakeven point is where price equals the minimum average total cost (Figure 8.9). If the price is between average total cost and average variable cost, then the firm is not breaking even; it is losing money. However, it does not make sense to shut down if the price is above the shutdown point.

REVIEW

- Profits can be represented as a rectangle on the cost curve diagram. So can losses. The profit or loss rectangle is the difference between the revenue rectangle and the loss rectangle.
- At the breakeven point, profits are zero. When $P < AVC$, profits are maximized by shutting down.

Costs and Production: The Long Run

Thus far, we have focused our analysis on the short run. By definition, the short run is the period of time during which it is not possible for firms to adjust certain inputs to production. But what happens in the long run, when it *is* possible for firms to make such adjustments? For example, what happens to On-the-Move when it opens new terminals or takes out a lease on a fleet of new trucks? To answer this question, we need to show how the firm can adjust its fixed costs as well as its variable costs in the long run. All costs can be adjusted in the long run.

The Effect of Capital Expansion on Total Costs

First, consider what happens to fixed costs when the firm increases its capital. Suppose On-the-Move increases the size of its fleet from 4 trucks to 8 trucks and raises the number of terminals from 2 to 4. Then its fixed costs would increase because more rent would have to be paid for 4 terminals and 8 trucks than for 2 terminals and 4 trucks. To obtain the increase in fixed costs, we need to use the price of capital. Again, suppose trucks cost $25 per day and a terminal costs $100 per day. Then 4 trucks and 2 terminals would cost $300 and 8 trucks and 4 terminals would cost $600. Fixed costs would rise from $300 to $600.

Second, consider what happens to variable costs when the firm increases its capital. An increase in capital increases the amount that each additional worker can produce. For example, according to Table 8.2, 166 worker-hours were required for On-the-Move to move 10 pianos when there were 4 trucks and 2 terminals. With more capital (8 trucks and 4 terminals), it will take fewer hours of work to move the pianos. Assume, for example, that it takes only 120 worker-hours to deliver 10 pianos. In this

scenario, with the wage equal to $10 per hour, the variable cost of moving 10 pianos falls from $1,660 to $1,200. In other words, variable costs decline as the firm expands its capital.

Now consider total costs. With fixed costs larger and variable costs smaller as a result of the increase in capital, what is the effect on total costs, which are the sum of fixed costs and variable costs? After the expansion, total costs will be higher at very low levels of output, where fixed costs dominate, but will be lower at high levels of output, where variable costs dominate. Figure 8.11 illustrates this using the total cost curve. Figure 8.11 is essentially the same as Figure 8.2 except that the green curves show the old costs before the expansion of capital and the purple curves show the new costs with the additional capital. The diagram shows that the new fixed costs are higher and the new variable costs are lower. The new total cost curve (TC_2) is twisted relative to the old total cost curve (TC_1). The new total cost curve is above the old total cost curve at low levels of output and below the old total cost curve at high levels of output.

Table 8.3 provides the numerical information about the costs that appear in Figure 8.11. To see the effect of the firm's expansion on its costs, compare the fixed costs, variable costs, and total costs in Table 8.3 with those in Table 8.1. Observe that fixed costs are higher: $600 rather than $300. Variable costs are lower throughout the range of production. As a result, total costs are higher in Table 8.3 than in Table 8.1 for production of less than 8 units, and lower in Table 8.3 than in Table 8.1 for production of 8 units or more.

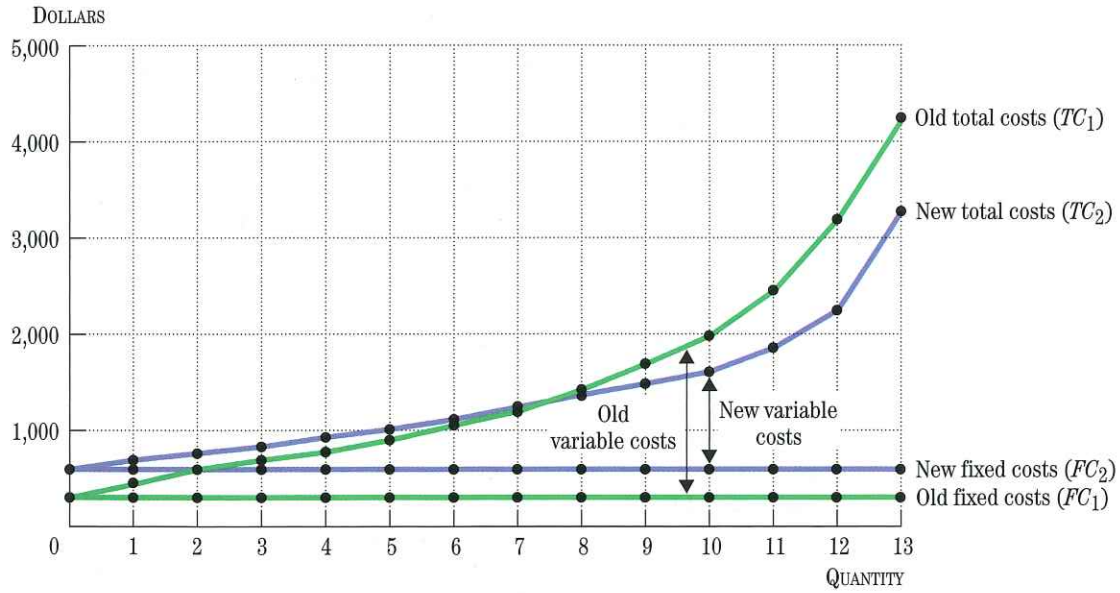


Figure 8.11
Shifts in Total Costs as a Firm Increases Its Capital in the Long Run

When a firm increases its capital, its fixed costs increase; as shown in the diagram, fixed costs rise from FC_1 to FC_2. However, its variable costs decrease, which is also shown. Thus, the new total cost curve (TC_2) will be above the old total cost curve (TC_1) for low-level output and below the old total cost curve (TC_1) for high-level output.

Table 8.3
Costs with More Capital (Compared with Table 8.1, fixed costs are higher and variable costs are lower in this table because capital is higher than in Table 8.1. Costs are measured in dollars.)

Quantity	Total Costs	Fixed Costs	Variable Costs	Average Total Cost
0	600	600	0	—
1	690	600	90	690
2	770	600	170	385
3	840	600	240	280
4	920	600	320	230
5	1,010	600	410	202
6	1,110	600	510	185
7	1,220	600	620	174
8	1,340	600	740	168
9	1,470	600	870	163
10	1,610	600	1,010	161
11	1,880	600	1,280	171
12	2,300	600	1,700	192

Effects of a Capital Expansion on Average Total Cost

Our analysis of the effects of the firm's capital expansion on total costs can be used to derive the effects on average total cost. Remember, average total cost (ATC) is total costs (TC) divided by the quantity (Q). Thus, if total costs increase at a given quantity of output, so will average total cost. And if total costs decrease at a given level of output, so will average total cost.

This is illustrated in Figure 8.12. An average total cost curve, labeled ATC_1, corresponding to average total cost in Table 8.1 is plotted. Another average total cost curve, labeled ATC_2, corresponding to average total cost in Table 8.3 is also plotted. The new average total cost curve (ATC_2) is above the old average total cost curve (ATC_1) at low levels of output and below the old average total cost curve (ATC_1) at higher levels of output. Average total cost is higher for production of less than 8 units and lower for production of 8 units or more. This is precisely what is shown for total costs in Figure 8.11.

The Long-Run *ATC* Curve

Now that we have seen what happens at On-the-Move when capital is expanded by a certain amount, we can see what happens when capital increases by even larger amounts. For example, suppose On-the-Move expands throughout Houston and even beyond Houston by expanding the size of its fleet of trucks and the number of terminals to park and service the trucks.

Figure 8.13 shows four different average total cost curves. Each of the average total cost curves corresponds to increased capital expansion at On-the-Move. The first two of these, ATC_1 and ATC_2, are the average total cost curves from Figure 8.12. Note that the second curve (ATC_2) is above the first (ATC_1) at low levels of output and below the first at high levels of output. The third and fourth curves are for even more trucks and terminals. The third average total cost curve (ATC_3) is above the second

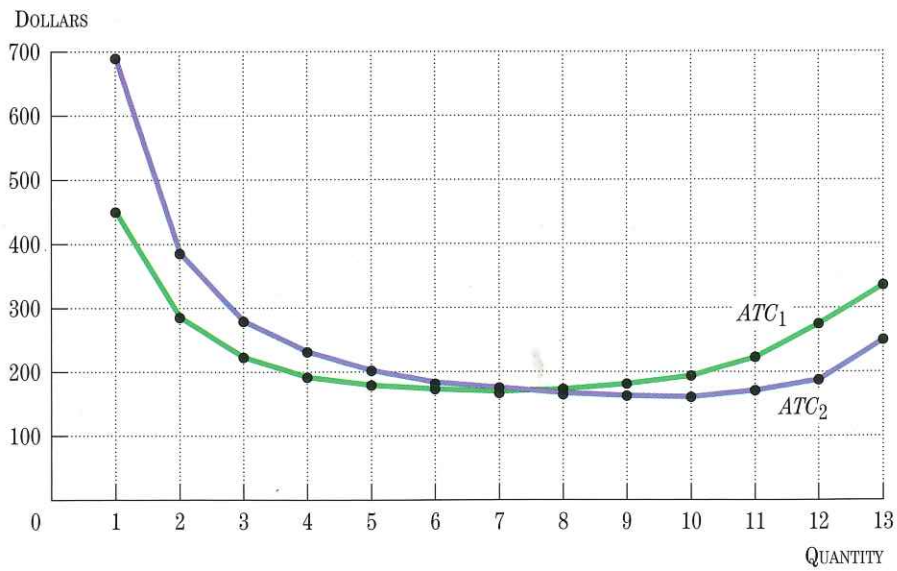


Figure 8.12
Shifts in Average Total Cost Curves When a Firm Expands Its Capital
The effects on average total cost follow directly from the effects on total costs in Figure 8.11. Here ATC_1 is the average total cost curve with a lower amount of capital, and ATC_2 is the average total cost curve with a higher amount of capital. To the left, at lower levels of output, higher fixed costs raise average total cost; to the right, at higher levels of output, lower variable costs tend to lower average total cost.

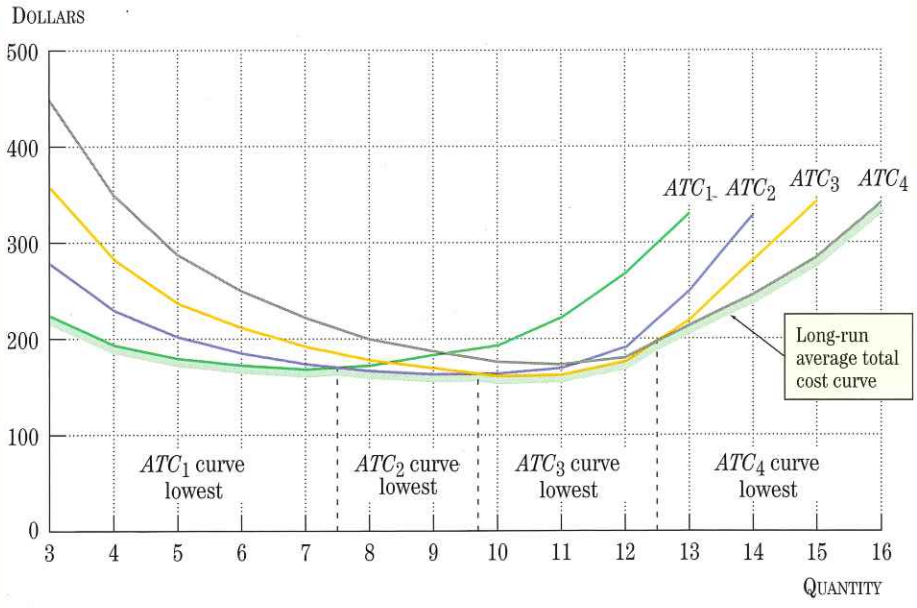


Figure 8.13
Long-Run versus Short-Run Average Total Cost
In the short run, it is not possible to change certain inputs, like the size of the factory or the number of machines. In the long run, these can be changed. For example, in the long run, On-the-Move can build more terminals around town. This means that the *ATC* curve shifts. The diagram shows four different *ATC* curves for On-the-Move; each new *ATC* curve represents more terminals, buildings, and machines than the *ATC* curve to its left. The long-run average total cost curve is shown by the thicker light green line.

(ATC_2) at lower levels of output and below the second (ATC_2) at higher levels of output. The same is true of the fourth compared to the third.

The thick light green curve tracing out the bottoms of the four average total cost curves gives the lowest average total cost at any quantity produced. For example, at 11 pianos moved, the lowest average total cost is $164. This occurs on the average total cost curve ATC_3. This thick line tells us what average total cost is when the firm can expand (or contract) its capital; in other words, this is the average total cost curve for the long run. For this reason, we call the curve that traces out the points on the lowest average total cost curves the **long-run average total cost curve.** The other average total cost curve that we have been discussing is called the *short-run average total cost curve*, or simply the average total cost curve.

long-run average total cost curve: the curve that traces out the short-run average total cost curves, showing the lowest average total cost for each quantity produced as the firm expands in the long run.

The long-run average total cost curve is one way in which economists study the behavior of a firm over time. Frequently economists will simply draw a generic long-run average total cost curve without the short-run curves. Whether the long-run average total cost curve slopes up or down, and over what range, is crucial for understanding the nature of a firm, the industry in which it operates, and the role of government.

The lack of smoothness in the long-run average total cost curve may seem strange. It occurs in Figure 8.13 because we have drawn only four short-run average total cost curves. If it is possible to expand capital in smaller amounts, then the curve will look smoother. For example, between the first and second short-run average total cost curves (ATC_1 and ATC_2) in Figure 8.13, there might be an average total cost curve for 6 trucks and 3 terminals. When we put in more and more short-run average total cost curves, the long-run average total cost curve gets smoother and smoother. But it still simply traces out the points of lowest cost for each level of output.

Capital Expansion and Production in the Long Run

How does a firm like On-the-Move decide whether to expand or contract its capital in the long run? How much does it produce in the long run? The decision is similar to the short-run decision about whether to hire more workers to move more pianos. Again, the firm sets the quantity produced to maximize profits. But now the quantity produced and profits can be affected by changes in capital as well as labor. At any level of capital and labor input, we can compute profits. For each level of capital, there is a short-run average total cost curve. Profits can be computed from this average total cost curve as described in Figure 8.7. Hence, for any level of capital, profits can be computed by the firm.

If the firm can increase its profits by expanding its capital and its output, then we predict that it will do so. If we find that the firm can increase its profits by reducing its capital and its output, then we predict that it will do so. In other words, the firm adjusts the amount of capital to maximize profits.

The Mix of Capital and Labor

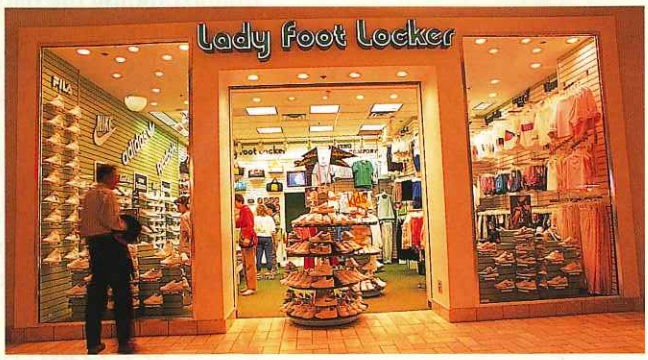
In the long run, the firm adjusts both its capital and its labor. Both inputs are variable in the long run. What determines the mix of labor and capital when both are variable? The relative price of labor compared to capital will be the deciding factor.

In deriving the cost curves for On-the-Move, we assumed that the cost of labor was $10 per hour and that the cost of capital, consisting of trucks and terminals, was $25 a day for trucks and $100 a day for terminals. If the cost of labor was higher, say, $20 per hour, then On-the-Move would have the incentive to rent more trucks rather than hire more workers, at least to the extent that this was feasible.

Economics in Action

Expanding (and Shrinking) a Firm over Time

The first general-merchandise "five and dime" store was opened by Woolworth in 1879. By building new stores and merging with other firms over time, Woolworth expanded greatly in size. By 1919 it had 1,081 stores in the United States, France, England, and Germany. Through the development of discount stores (Woolco) and specialty stores (including the Lady Foot Locker shown here), this firm continued to expand. In 1999 the company dropped the name Woolworth and became Venator. The newly named firm sold many of its non-footwear businesses and eventually was renamed Foot Locker, Inc., in November 2001. By 2005, Foot Locker had over 4,000 specialty stores in Australia, Europe, New Zealand, and North America.



However, if the cost of capital rose relative to that of labor, then the firm would have the incentive to hire more workers. In general, the firm will use more capital relative to labor if the cost of capital declines relative to that of labor. And conversely, the firm will use less capital relative to labor if the cost of capital rises relative to that of labor.

REVIEW

- In the long run, the firm can expand by increasing its capital. Fixed costs increase and variable costs decline at each level of production as the firm expands its capital. Thus, total costs, and average total cost, are higher at low levels of production and lower at high levels of production.
- The long-run average total cost curve traces out the points on the lowest short-run average total cost curve for each level of production.

Economies of Scale

The long-run average total cost curve describes a situation in which the firm can expand all its inputs—both its capital and its labor. When all inputs increase, we say that the *scale* of the firm increases. For example, if the number of workers at the firm

doubles, the number of trucks doubles, the number of terminals doubles, and so on, then we say that the scale of the firm doubles. Thus, the long-run average total cost curve describes what happens to a firm's average total cost when its scale increases. There is some specialized terminology about different shapes of the long-run average total cost curve.

economies of scale: a situation in which long-run average total cost declines as the output of a firm increases.

diseconomies of scale: a situation in which long-run average total cost increases as the output of a firm increases.

constant returns to scale: a situation in which long-run average total cost is constant as the output of a firm changes.

We say that there are **economies of scale,** or *increasing returns to scale,* if long-run average total cost falls as the scale of the firm increases. We say that there are **diseconomies of scale,** or *decreasing returns to scale,* if long-run average total cost rises as the scale of the firm increases. The situation in the middle, where long-run average total cost neither rises nor falls, is called **constant returns to scale.** Figure 8.14 illustrates these three possible shapes for the long-run average total cost curve.

Determining Whether a Firm Has Economies or Diseconomies of Scale

Whether there are increasing, decreasing, or constant returns to scale depends on the type of firm and the type of product. Consider a firm like On-the-Move. One can imagine that there would be economies of scale as the firm expanded the number of terminals around the city of Houston; with more terminals, trucks could be serviced at many different locations and would not have to be driven so far at the end of the day or towed so far in the event of a breakdown. With a larger work force, On-the-Move could have workers who *specialize* in moving different types of pianos or who specialize in servicing different parts of the trucks. Some might specialize in moves to high-rise buildings. In other words, as the scale of a firm increases, the work can be divided into different tasks and some members of the labor force can specialize in each task.

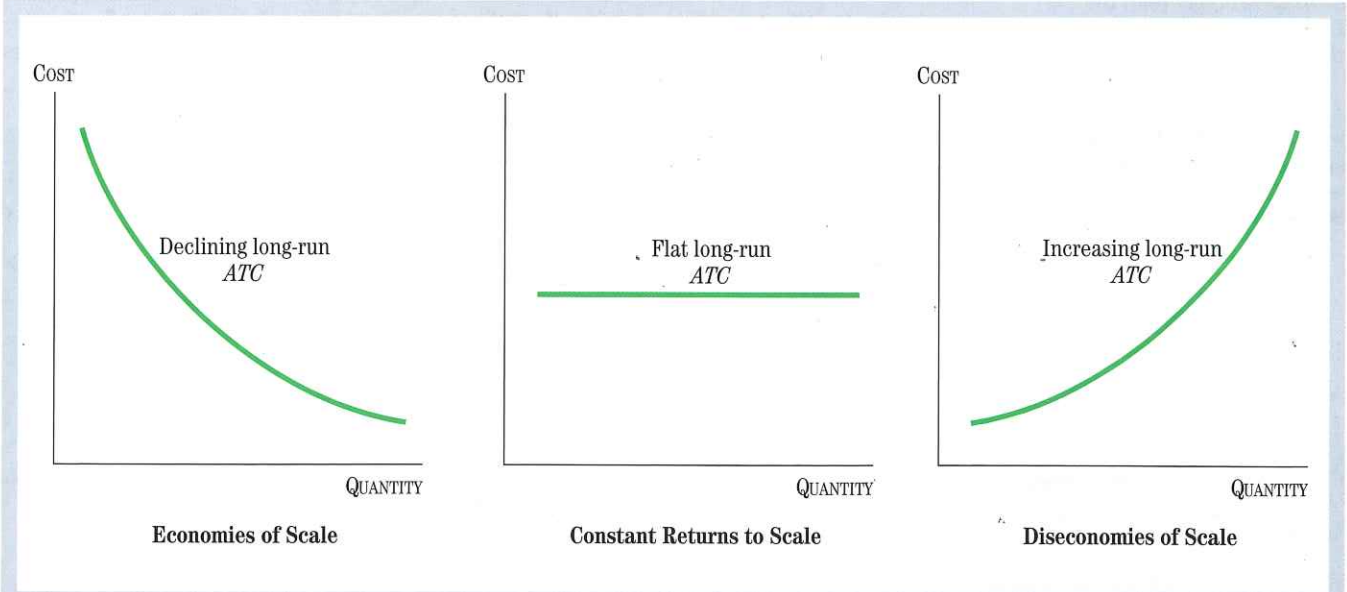


Figure 8.14 Economies and Diseconomies of Scale

If the long-run average total cost curve slopes downward, we say there are economies of scale. If the long-run average total cost curve slopes upward, we say there are diseconomies of scale. If the long-run average total cost curve is flat, we say there are constant returns to scale, as shown in the middle panel.

Gulliver and Economies of Scale
Like a firm expanding all *its inputs—capital and labor—Gulliver found that he was bigger in* all *dimensions—arms and legs and head—than the Lilliputians. When economists consider economies of scale, they think of all the firm's inputs increasing, not just one. But when they consider diminishing returns to labor, they think of only one input (labor) increasing, which would be like Gulliver's finding that only his arms were bigger than the Lilliputians', with everything else the same size.*

Is there a limit to economies of scale? What about expanding beyond Houston to Galveston, Dallas, Tulsa, Mexico City, or even Lima, Peru? In the case of piano moving, returns to scale would probably begin to decline at some point. The extra administrative costs of organizing a large interstate or worldwide piano-moving firm would probably raise average total cost. Thus, one could imagine that the long-run average total cost curve for On-the-Move would first decline and then increase as the firm grows in size.

Although no two firms are alike, the long-run average total cost curve for most firms probably declines at low levels of output, then remains flat, and finally begins to increase at high levels of output. As a firm gets very large, administrative expenses, as well as coordination and incentive problems, will begin to raise average total cost. The smallest scale of production for which long-run average total cost is at a minimum is called the **minimum efficient scale.** A typical long-run average cost curve and its minimum efficient scale are shown in Figure 8.15.

Mergers and Economies of Scope

An increase or decrease in the scale of a firm through capital expansion or contraction—as described in the previous two sections—is one kind of change in the firm over time. Firms can also change over time in other ways. They can grow through mergers between one firm and another firm. If the product lines of the two firms are similar, then such mergers may be a way to reduce costs. That is one reason why large oil companies, such Exxon and Mobil, merged in the 1990s. Mergers are also a common way for firms to combine different skills or resources to develop new products. For example, America Online, an Internet firm, and Time Warner, a large firm producing movies, magazines, and CDs, merged in 2001. By bringing together distribution resources with content resources, this merger was intended to widen the scope of both firms and help them develop new products. The results did not turn out as expected, however, and in 2003 Time Warner dropped AOL from its corporate name in an effort to demonstrate that the company still valued its "core" assets—magazines, books, cable (HBO, CNN), and movies. Combining different types of firms

minimum efficient scale: the smallest scale of production for which long-run average total cost is at a minimum.

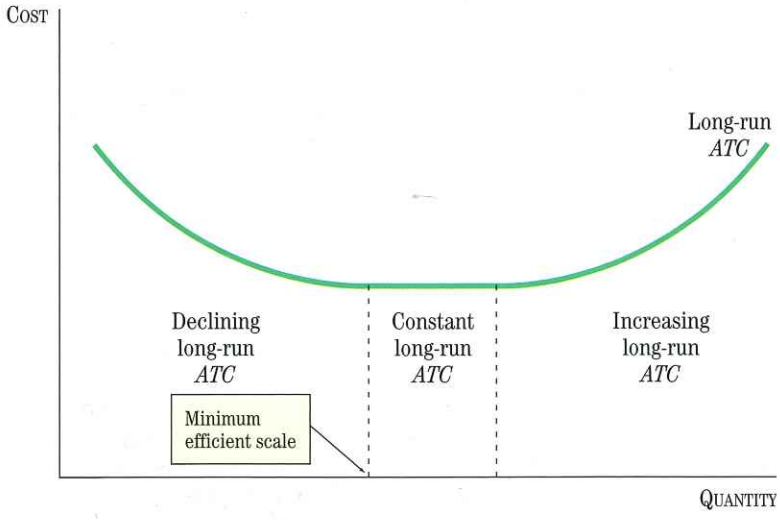


Figure 8.15
Typical Shape of the Long-Run Average Total Cost Curve
For many types of firms, the long-run *ATC* curve slopes down at low levels of output, then reaches a flat area, and finally begins to slope up at high levels. The minimum efficient scale is shown.

Table 8.4
Recent Big Mergers

Names Before Merger		Name After Merger	Primary Industry
Exxon	Mobil	ExxonMobil	Oil
BP	Amoco	BP Amoco (now BP)	Oil
J.P. Morgan	Chase Manhattan	J.P. Morgan Chase	Banking
Bank of America	Fleet Boston Financial	Bank of America Corp.	Banking
Chrysler	Daimler-Benz	DaimlerChrysler	Motor vehicles
WorldCom	MCI Communications	MCI WorldCom (now WorldCom)	Telecommunications
Walt Disney	Capital Cities/ABC	Walt Disney	Entertainment
America Online	Netscape	America Online	Internet
America Online	Time Warner	AOL Time Warner	Internet + entertainment

Most of these mergers are in the same industry and illustrate economies of scale except for the last, which may be due to expected economies of scope.

to reduce costs or create new products is sometimes called *economies of scope*. Table 8.4 lists some of the big mergers of recent years. Observe that most of those listed involve similar products.

When two firms merge into one firm, each of the original firms may become a division within the new firm. A merger of one firm with another means that the coordination of production moves from the market to within the organization. For example, the merger of the U.S. firm Chrysler with the German firm Daimler-Benz meant that the cars were produced in two divisions of the same firm.

REVIEW

- Economies of scale are nothing more than a downward-sloping long-run average total cost curve. Economies of scale may occur because of the specialization that the division of labor in larger firms permits.
- Although economies of scale probably exist over some regions of production, the evidence indicates that as firms grow very large, diseconomies of scale set in.
- Firms sometimes expand by merging with other firms. Many big mergers occurred in the 1990s.

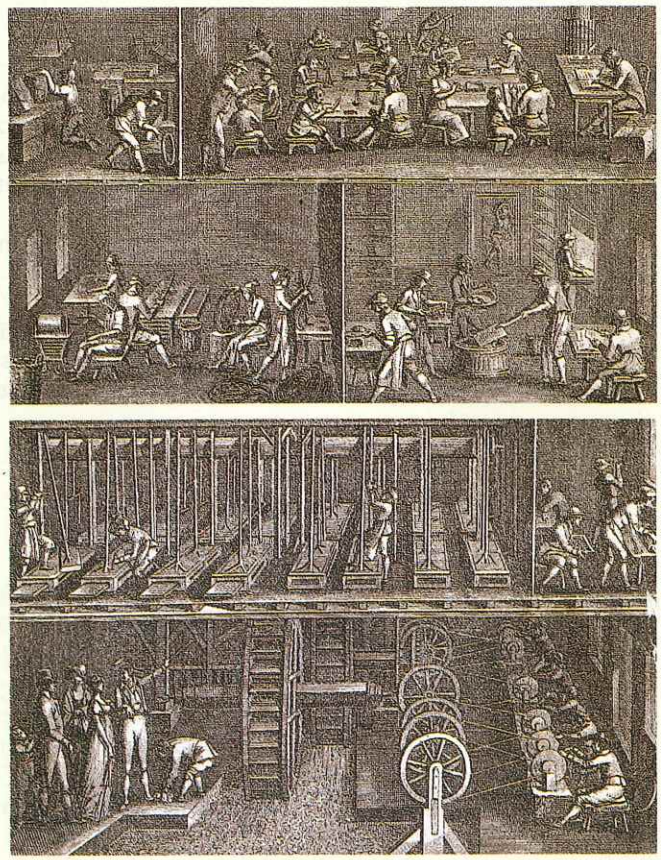
Conclusion

In this chapter we have developed a model for studying why firms shut down, expand, or contract. This long-term analysis of changes at firms over time is an extension of the analysis of a firm's short-run behavior in Chapter 6.

A centerpiece of the model is a graph (Figure 8.5) that shows the firm's average total cost and average variable cost. Using this graph, we can determine whether a firm will shut down in the short run. Using additional average total cost curves corresponding to alternative levels of capital, we can also determine whether the firm will expand or contract.

Economist Finds Economies of Scale at Pin Factory

One way to get an intuitive feel for economies of scale is to visit a factory and watch people in action. If you look carefully, you can actually see economies of scale. Here is a short but wonderfully vivid description of workers at a pin-making factory. An economist wrote the description after a visit to the factory, and it illustrates how economists think. The workers are producing pins from metal wire.



> One man draws out the wire, another straightens it, a third cuts it, a fourth points it, a fifth grinds it at the top for receiving the head; to make the head requires two or three distinct operations; to put it on is a peculiar business, to whiten the pins is another . . . [T]en men only . . . could make upwards of forty eight thousand pins in a day. Each person, therefore, . . . might be considered as making four thousand eight hundred pins in a day.
>
> But if they had each wrought separately and independently they certainly could not each of them have made twenty, perhaps, not one pin in a day.
>
> That is certainly not the two hundred and fortieth, perhaps not the four thousand eight hundredth part of what they are at present capable of producing.

This story of the pin factory is not new. In what century do you think the economist watched this firm in action? And who do you think the economist was? (*Hint:* It was written up in a book called *Wealth of Nations*.) Whatever the answer, the story is as relevant now as it was when the account was written.

A more modern example of economies of scale involves Wal-Mart, the retail company founded by Sam Walton in 1962, now selling $217.8 billion and employing over 1,300,000 employees. How did Wal-Mart grow to be the largest U.S. company in revenues in just forty years?

One of the keys to Wal-Mart's success is the use of economies of scale. With over 3,000 stores in the United States and over 1,000 additional stores in nine other countries, Wal-Mart uses central distribution points, a streamlined supply chain closely coordinated with its suppliers, lower inventories, centrally coordinated marketing, and its enormous buying power to lower its costs. Wal-Mart's focus on economies of scale, increased productivity, and cost reduction allowed it to charge lower prices and increase its market share from 9% in 1987 to 27% in 1995. As an example of Wal-Mart's advantage over other general-merchandise retailers, its productivity measured by real sales per employee was 36% higher than its closest competitor in 1999. As a matter of fact, the economies of scale of the retail industry—particularly Wal-Mart—contributed greatly to the increase in U.S. productivity in the 1990s.

What similarities and differences do you see between Wal-Mart and the pin factory? What aspects other than economies of scale would you consider when measuring Wal-Mart's success and its impact on the economy?

The economist looks at what each worker does, and estimates that production each day at this ten-person firm equals 48,000 pins, or 4,800 per worker.

Now the economist imagines production at a much smaller firm consisting of one worker. Daily production is estimated at between 1 and 20 pins per worker.

Thus *production per worker* at the small firm is only 1/240 or 1/4,800 as much as at the large firm. That is big economies of scale.

By looking at a firm's long-run average total cost curve (Figure 8.15), we can tell whether a firm has economies of scale. We will use the average total cost curve extensively in the next several chapters of this book.

KEY POINTS

1. Firms start up, expand, contract, or shut down when conditions in the economy change.
2. The short run and the long run are two broad categories into which economists categorize time periods. The short run is the period of time during which it is not possible for the firm to change all the inputs to production; only some inputs, such as labor, can be changed. The long run is the minimum period of time in which the firm can vary all inputs to production, including capital.
3. Average total cost, or cost per unit, is widely used by economists, accountants, and investors to assess a firm's cost behavior.
4. When the market price equals the minimum of average total cost, the firm breaks even. At higher prices, profits are positive. At lower prices, profits are negative.
5. When the market price equals the minimum of average variable cost, the firm is just at the point of shutting down. If the price is below average variable cost, the firm should shut down.
6. The long-run average total cost curve describes how a firm's costs behave when the firm expands its capital.
7. If long-run average total cost declines, then there are economies of scale. For many firms, there is a range over which the long-run average total cost curve is flat, and we say that there are constant returns to scale. When firms get very large, diseconomies of scale set in.
8. Firms can expand by merging with other firms. Such mergers are motivated by either economies of scale or economies of scope.

KEY TERMS

short run
long run
average total cost (*ATC*)
average variable cost (*AVC*)
average fixed cost (*AFC*)
average product of labor
breakeven point
shutdown point
long-run average total cost curve
economies of scale
diseconomies of scale
constant returns to scale
minimum efficient scale

QUESTIONS FOR REVIEW

1. What is the difference between average total cost and average variable cost?
2. Why does the marginal cost curve cut through the average total cost curve exactly at the minimum of the average total cost curve?
3. Why are total revenue, total costs, and profits given by areas of rectangles in the cost curve diagram?
4. What is the difference between the breakeven point and the shutdown point?
5. Why do average total cost curves shift as the firm expands, and how does the shift relate to economies of scale?
6. Why might a merger lower average total cost?
7. What is the minimum efficient scale of a firm?

PROBLEMS

1. Draw the typical average total cost, average variable cost, and marginal cost curves for a profit-maximizing, price-taking firm.
   a. Show the case where price equals average total cost.
   b. Show the rectangles that represent fixed costs and variable costs. What happens to the size of these areas as the market price increases? Show this in your diagram.
2. Consider the age of the people in a restaurant. Suppose the first person you notice is 40 years old and the second and third are 33 and 27, respectively.
   a. Graph the average and marginal age of the three people in the restaurant, placing age on the vertical axis and quantity of people on the horizontal axis in the same order in which you notice them.
   b. What do you notice about the relationship between marginal and average age?
   c. Suppose the fourth person entering the restaurant is 41 years old. What happens to the average age?

3. Fill out the table below, and then answer the following questions.

Q	*TC*	*FC*	*VC*	*ATC*	*AVC*	*MC*
0	8					
1	12					
2	14					
3	20					
4	30					
5	50					

   a. Suppose the firm is a price-taker. If the price is $15 per unit, will this firm be earning economic profits? How much? What quantity will it produce?
   b. What is the breakeven price? the shutdown price?

4. Suppose the firm in problem 3 can buy an additional machine that causes its minimum short-run average total cost to become $10. Does this expansion involve economies or diseconomies of scale? Sketch this situation in a diagram.

5. Plot the following data on quantity of production and long-run average total cost for a firm. Show the areas of economies and diseconomies of scale and constant returns to scale. What is the minimum efficient scale?

Quantity	Long-Run *ATC*
1	33
2	27
3	25
4	25
5	30
6	38
7	50

6. Suppose the average total cost curves for a firm for three different amounts of capital are as follows:

Quantity	ATC_1	ATC_2	ATC_3
1	40	50	60
2	30	35	40
3	20	25	30
4	30	15	25
5	40	30	20
6	50	40	30

   a. Plot the three average total cost curves in the same diagram.
   b. Determine the long-run average total cost curve and show it in the same diagram as in part (a).

7. Fill out the table below for a competitive firm that can sell its product for $13 a unit.
   a. What quantity will this firm produce? Why?
   b. At that quantity level, what profits or losses will this firm make?
   c. Is this market in long-run equilibrium? Why or why not?
   d. In the short run, at what price would this firm break even? At what price would the firm shut down? Explain briefly.

8. Suppose you paid $20 at the start of the term for a four-month subscription to the *Economist* magazine. Then, after one week, you find out that reading the magazine will not help you get a better grade in your economics course and that you do not enjoy reading it. Your parents advise you to read the magazine anyway because you paid for it. Is this good advice? Write a letter to your parents explaining the concept of sunk costs and relate it to their advice.

Problem 7

Quantity	Total Cost	Fixed Cost	Variable Cost	Average Total Cost	Average Variable Cost	Marginal Cost	Total Revenue	Marginal Revenue	Profit
0									
1				$27.00		$9.00			
2				$16.00					
3					$5.00				
4					$5.50				
5					$8.40				

Producer Theory with Isoquants

In this chapter we looked at how firms adjust their labor and capital inputs when the cost of these inputs changes over time. Here we give a graphical illustration of a firm's choice between labor and capital. The graphs are similar to the budget lines and indifference curves used to describe consumer choice in the appendix to Chapter 5. We use these graphs to show exactly how a firm's choice between labor and capital depends on the relative price of labor and capital.

Combining Capital and Labor

Consider an example of a firm with two inputs to production: capital and labor. Table 8A.1 shows the possible combinations of inputs available to the firm. For example, if the firm has 2 units of capital and uses 24 hours of labor, it can produce 3 units of output. The hypothetical numbers in Table 8A.1 could represent a wide variety of firms producing different types of products, but observe that we have chosen the units in the table to be the same as those for the firm shown in Table 8.2. (To make a comparison between Table 8A.1 and Table 8.2, you can think of a "unit" of capital as corresponding to 4 trucks and 2 terminals with a cost of $300; 2 units of capital is 8 trucks and 4 terminals at a cost of $600.) Table 8A.1 could refer to a firm with any type of capital (computers, machine tools, telephones, or pizza ovens). To allow for all these possibilities, we refer to capital as a "unit" of capital.

The information in Table 8A.1 can be represented graphically, as shown in Figure 8A.1. Each column is plotted with labor input on the horizontal axis and the quantity produced on the vertical axis. Each column represents the production function for a given level of capital. Note that higher levels of capital increase the amount that can be produced with a given amount of labor. In

Table 8A.1
Production with Four Levels of Capital

	Labor Input (hours)			
Quantity Produced	*With 1 Unit of Capital*	*With 2 Units of Capital*	*With 3 Units of Capital*	*With 4 Units of Capital*
0	0	0	0	0
1	15	9	6	5
2	27	17	12	10
3	37	24	17	13
4	48	32	22	18
5	60	41	29	23
6	74	51	36	29
7	90	62	43	35
8	109	74	52	41
9	134	87	61	49
10	166	101	71	57
11	216	128	90	72
12	290	170	119	95
13	400	270	189	151
14	—	400	300	220
15	—	—	425	300
16	—	—	—	430

Note: The column showing labor input with 1 unit of capital corresponds to the production function for On-the-Move discussed in Chapter 8 (see Table 8.2). The omitted entries in the table represent quantities of production that cannot be achieved without more capital.

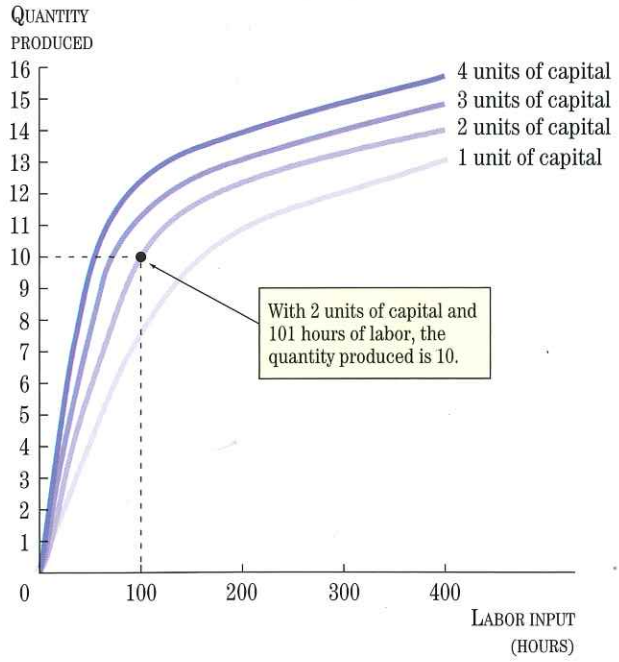


Figure 8A.1
The Production Function with Four Levels of Capital
As the amount of labor input increases, so does the amount of output. Each curve corresponds to a different level of capital. Higher curves represent higher levels of capital. The points on these four curves are obtained from the four columns of Table 8A.1.

8A.1. The same information appears in a different and convenient way.

Isocost Lines

A firm's total costs can also be shown on a diagram like Figure 8A.2. In considering the choice between capital and labor, the firm needs to consider the price of both. Suppose that labor costs \$10 per hour and capital costs \$300 per unit. Then if the firm uses 1 unit of capital and 150 hours of labor, its total costs will be $1 \times \$300 + 150 \times \$10 = \$1{,}800$. For the same total costs, the firm can pay for other combinations of labor and capital. For example, 2 units of capital and 120 hours of labor also cost \$1,800. Other combinations are as follows:

Hours of Labor	Units of Capital	Total Costs
180	0	$180 \times \$10 + 0 \times \$300 = \$1{,}800$
150	1	$150 \times \$10 + 1 \times \$300 = \$1{,}800$
120	2	$120 \times \$10 + 2 \times \$300 = \$1{,}800$
90	3	$90 \times \$10 + 3 \times \$300 = \$1{,}800$
60	4	$60 \times \$10 + 4 \times \$300 = \$1{,}800$
30	5	$30 \times \$10 + 5 \times \$300 = \$1{,}800$
0	6	$0 \times \$10 + 6 \times \$300 = \$1{,}800$

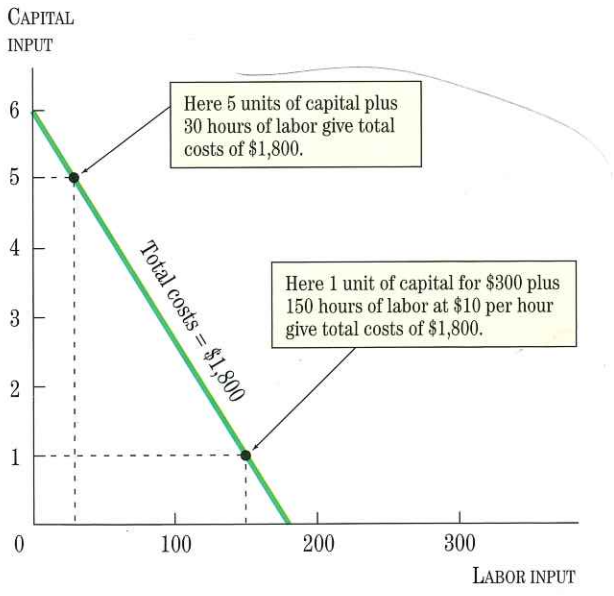


Figure 8A.3
An Isocost Line
Each isocost line shows all the combinations of labor and capital that give the same total costs. In this case, the price of capital is \$300 per unit and the price of labor is \$10 per hour. Total costs are \$1,800. For example, if 1 unit of capital and 150 hours of labor are employed, total costs are \$1,800 = (1 × \$300) + (150 × \$10).

In other words, the \$1,800 can be spent on any of these combinations of labor and capital. With \$1,800, the firm can use 6 units of capital, but that would not permit the firm to hire any workers.

These different combinations of labor and capital that have total costs of \$1,800 are plotted in Figure 8A.3. Each combination of labor and capital in the table is plotted, and the points are connected by a line. The line is called an **isocost line.** An isocost line shows the combinations of capital and labor that have the same total costs.

The position of the isocost line depends on the amount of total costs. Higher total costs are represented by higher isocost lines. This is shown in Figure 8A.4. Observe that the isocost line for total costs of \$2,100 is above the one for total costs of \$1,800.

The slope of the isocost line depends on the ratio of the price of labor to the price of capital. In particular, the slope equals −1 times the ratio of the price of labor to the price of capital. This is illustrated in Figure 8A.5 for the case where total costs equal \$1,800. If the price of labor falls from \$10 to \$6, then the isocost line gets flatter. Thus, if the hourly wage were \$6 instead of \$10, the firm

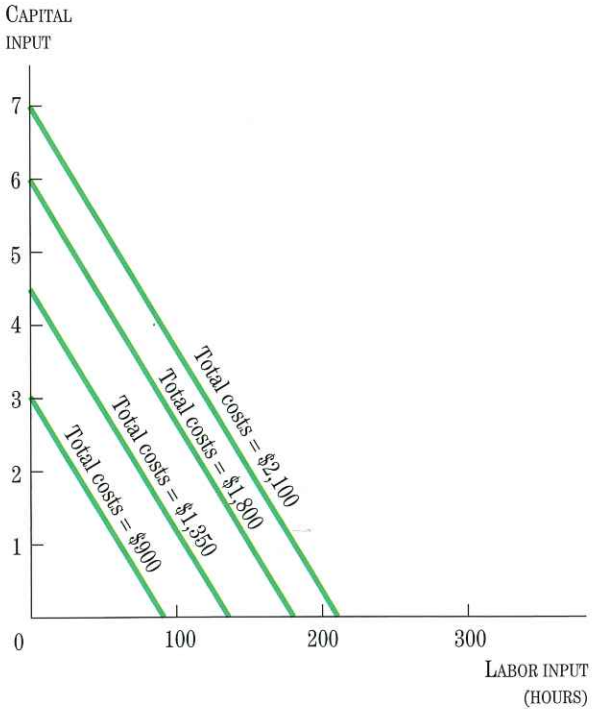


Figure 8A.4
Several Isocost Lines with Different Total Costs
Isocost lines with higher total costs are above and to the right of those with lower total costs. All the isocost lines in this diagram have a capital cost of \$300 per unit and a labor cost of \$10 per hour.

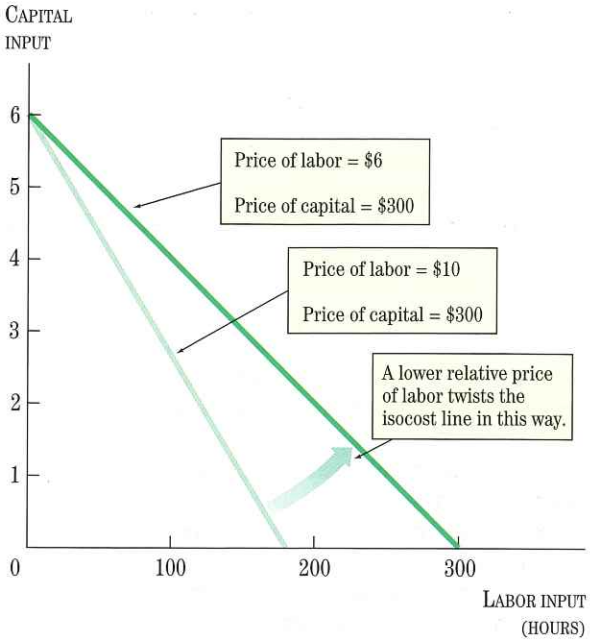


Figure 8A.5
Effect of a Change in Relative Prices on the Isocost Line
When the price of labor falls relative to the price of capital, the isocost line gets flatter, as in this diagram. In this case, the price of labor falls from $10 per hour to $6 per hour while the price of capital remains at $300 per unit. Total costs remain equal to $1,800 in this case.

would be able to pay for 250 hours of work and 1 unit of capital, compared with only 150 hours and 1 unit of capital, and still have total costs of $1,800. Thus, as the price of labor (the wage) falls relative to the price of capital, the isocost line gets flatter.

Minimizing Costs for a Given Quantity

The isoquant and isocost lines can be used to determine the least-cost combination of capital and labor for any given quantity of production. Figure 8A.6 shows how. In Figure 8A.6 we show three isocost lines, along with an isoquant representing 11 units of output. For the isocost lines, the price of labor is $10 and the price of capital is $300. The point where the isocost line just touches the isoquant is a *tangency point*. It is labeled *A*.

Point *A* is where the firm minimizes the cost of producing 11 units of output. To see this, suppose you are at point *A* and you move to the left and up along the same isoquant to point *B*. This means that the firm increases capital and decreases hours of labor, keeping the quantity produced constant at 11 units; that is, the firm substitutes capital for labor. But such a substitution increases the firm's costs, as shown in the figure. The payment for the extra capital will be greater than the saving from the reduced labor. Thus, moving along the isoquant from *A* to *B* would increase the total costs to the firm.

A similar reasoning applies to moving from point *A* to point *C*. The firm uses fewer labor hours and less capital at point *C*, so that total costs are lower than at point *A*. But at point *C* the firm does not have enough inputs to produce 11 units of output. Thus, point *A* is the lowest-cost point at which the firm can produce 11 units of output. It is the point at which the lowest isocost line is touching the isoquant.

The Cost Minimization Rule

The rate of technical substitution of capital for labor and the ratio of the price of labor to the price of capital are equal at point *A*, because the slopes of the isoquant and of the isocost line are equal at point *A*. If the firm is minimizing its costs, then the rate of technical substitution must equal the input price ratio. The equality between the rate of technical substitution and the input price ratio is called the *cost minimization rule.*

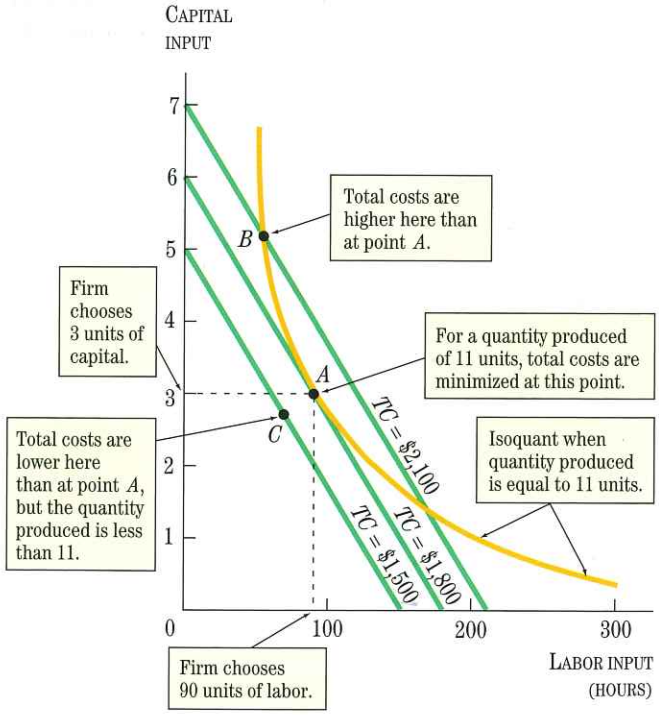


Figure 8A.6
Choosing Capital and Labor to Minimize Total Costs
The diagram illustrates how a firm chooses a mix of labor and capital to minimize total costs for a given level of output. Here the given level of output is 11 units, as shown by the single isoquant. Total costs are minimized by choosing the combination of labor and capital given by the tangency (point *A*) between the isocost line and the isoquant. For any other point on the isoquant, the quantity would be the same but total costs would be higher.

Observe that isoquants are analogous to the indifference curves and the isocost lines are analogous to the budget line described in the appendix to Chapter 5. The cost minimization rule for a firm is much like the utility maximizing rule for a consumer.

A Change in the Relative Price of Labor

Now we show how isoquants and isocost lines can be used to predict how a firm will adjust its mix of inputs when there is a change in input prices. For example, suppose that the hourly wage falls from \$10 to \$6 and the price of capital rises from \$300 to \$600. That is, labor becomes cheaper relative to capital. Originally, the ratio of the price of labor to capital was 10/300 = .033; now it is 6/600 = .010. This is a big reduction, and we would expect the firm to adjust by changing capital and labor input. Figure 8A.7 shows how the firm would adjust the mix of capital and labor for a given quantity of output. Figure 8A.7 keeps the isoquant fixed but includes a new isocost line that reflects the lower relative price of labor and is tangent to the isoquant. Since the new isocost line is flatter, the point of tangency with the given isoquant no longer occurs at point *A*, where 3 units of capital are combined with 90 hours of labor. Now tangency occurs at point *D*, where there is a combination of 2 units of capital and 130 hours of labor. In other words, the firm has substituted labor for capital when the relative price of labor fell. At the new point *D*, the firm would use 1 less unit of capital and 40 more hours of labor.

In summary, common sense tells us that the firm will hire more labor and use less capital when the price of labor falls relative to the price of capital. The isoquants and isocost lines confirm this and tell us by exactly how much.

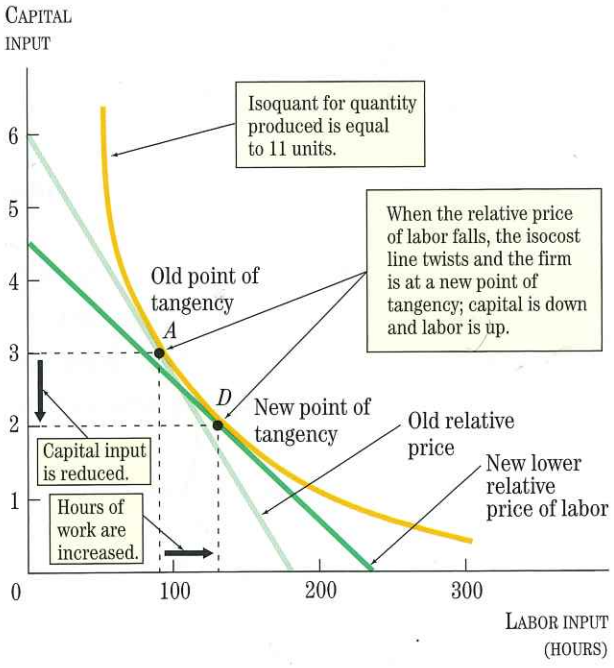


Figure 8A.7
Effect of a Lower Price of Labor Relative to Capital
The dark green isocost line has a lower price of labor relative to capital than the light green line. Hence, the amount of capital used by the firm decreases from 3 units to 2 units, and the amount of labor rises from 90 hours to 130 hours.

Key Terms and Definitions

isoquant: a curve showing all the possible combinations of two inputs that yield the same quantity of output.

rate of technical substitution: the rate at which one input must be substituted for another input to maintain the same production; it is the slope of the isoquant.

isocost line: a line showing the combinations of two inputs that result in the same total costs.

Questions for Review

1. Why does the isoquant slope downward?
2. Why does the isocost line slope downward?
3. What determines the slope of the isocost line?
4. Why does the firm minimize cost for a given level of output by choosing capital and labor at the point where the isocost line is tangent to the isoquant?

Problems

1. Graph the isocost line associated with a wage of \$10 per hour and a price of capital of \$50 for total costs of \$200, \$240, and \$300. Suppose the wage rises to \$15 and the price of capital stays at \$50. Show how the isocost line moves if total costs are \$300.
2. Sketch a typical isocost line and isoquant where the firm has chosen the combination of capital and labor that minimizes total costs for a given quantity of output. Now suppose the price of capital rises and the wage does not change. What must the firm do to maintain the same level of output as before the increase in the price of capital and still minimize costs? Will it substitute away from capital?
3. Draw a diagram with an isocost line and an isoquant next to a diagram with a budget line and an indifference curve from the appendix to Chapter 5. List the similarities and differences. How are the isocost and budget lines analogous to each other? How are the isoquant and the indifference curve analogous to each other? What is the importance of the tangency point in each case?