CHAPTER

6

Production and Business Organization

The business of America is business.

Calvin Coolidge

Before we can eat our daily bread, someone must bake it. Similarly, the economy's ability to build cars, generate electricity, write computer programs, and deliver the multitude of goods and services that are in our gross domestic product depends upon our productive capacity. Productive capacity is determined by the size and quality of the labor force, by the quantity and quality of the capital stock, by the nation's technical knowledge along with the ability to use that knowledge, and by the nature of public and private institutions. Why are living standards high in North America? Low in tropical Africa? For answers, we should look to how well the machine of production is running.

Our goal is to understand how market forces determine the supply of goods and services. Over the next three chapters we will lay out the essential concepts of production, cost, and supply and show how they are linked. We first explore the fundamentals of production theory, showing how firms transform inputs into desirable outputs. Production theory also helps us understand why productivity and living standards have risen over time and how firms manage their internal activities.

A. THEORY OF PRODUCTION AND MARGINAL PRODUCTS

BASIC CONCEPTS

A modern economy has an enormously varied set of productive activities. A farm takes fertilizer, seed, land, and labor and turns them into wheat or corn. Modern factories take inputs such as energy, raw materials, computerized machinery, and labor and use them to produce tractors, DVDs, or tubes of toothpaste. An airline takes airplanes, fuel, labor, and computerized reservation systems and provides passengers with the ability to travel quickly through its network of routes.

The Production Function

We have spoken of inputs like land and labor and outputs like wheat and toothpaste. But if you have a fixed amount of inputs, how much output can you get? On any day, given the available technical knowledge, land, machinery, and so on, only a certain quantity of tractors or toothpaste can be obtained from a given amount of labor. The relationship between the amount of input required and the amount of output that can be obtained is called the *production function.*

The **production function** specifies the maximum output that can be produced with a given quantity of inputs. It is defined for a given state of engineering and technical knowledge.

An important example is the production function for generating electricity. Visualize it as a book with technical specifications for different kinds of plants. One page is for gas turbines, showing their inputs (initial capital cost, fuel consumption, and the amount of labor needed to run the turbine) and their outputs (amount of electricity generated). The next page shows inputs and outputs of coal-fired generating plants. Yet other pages describe nuclear power plants, solar power stations, and so forth. Taken together, they constitute the production function for electricity generation.

Note that our definition assumes that firms always strive to produce efficiently. In other words, they always attempt to produce the maximum level of output for a given dose of inputs.

Consider the humble task of ditchdigging. Outside our windows in America, we see a large and expensive tractor, driven by one person with another to supervise. This team can easily dig a trench 5 feet deep and 50 feet long in 2 hours. When we visit Africa, we see 50 laborers armed only with picks. The same trench might take an entire day. These two techniques—one capital-intensive and the other labor-intensive—are part of the production function for ditchdigging.

There are literally millions of different production functions—one for each and every product or service. Most of them are not written down but are in people's minds. In areas of the economy where technology is changing rapidly, like computer software and biotechnology, production functions may become obsolete soon after they are used. And some, like the blueprints of a medical laboratory or cliff house, are specially designed for a specific location and purpose and would be useless anywhere else. Nevertheless, the concept of a production function is a useful way of describing the productive capabilities of a firm.

Total, Average, and Marginal Product

Starting with a firm's production function, we can calculate three important production concepts: total, average, and marginal product. We begin by computing the total physical product, or **total product,** which designates the total amount of output produced, in physical units such as bushels of wheat or number of sneakers. Figure 6-1(*a*) on page 109 and column (2) of Table 6-1 on page 110 illustrate the concept of total product. For this example, they show how total product responds as the amount of labor applied is increased. The total product starts at zero for zero labor and then increases as additional units of labor are applied, reaching a maximum of 3900 units when 5 units of labor are used.

Once we know the total product, it is easy to derive an equally important concept, the marginal product. Recall that the term "marginal" means "extra."

The **marginal product** of an input is the extra output produced by 1 additional unit of that input while other inputs are held constant.

For example, assume that we are holding land, machinery, and all other inputs constant. Then labor's marginal product is the extra output obtained by adding 1 unit of labor. The third column of Table 6-1 calculates the marginal product. The marginal product of labor starts at 2000 for the first unit of labor and then falls to only 100 units for the fifth unit. Marginal product calculations such as this are crucial for understanding how wages and other factor prices are determined.

The final concept is the **average product,** which equals total output divided by total units of input. The fourth column of Table 6-1 shows the average product of labor as 2000 units per worker with one worker, 1500 units per worker with two workers, and so forth. In this example, average product falls through the entire range of increasing labor input.

Figure 6-1 plots the total and marginal products from Table 6-1. Study this figure to make sure you understand that the blocks of marginal products in (*b*) are related to the changes in the total product curve in (*a*).

The Law of Diminishing Returns

Using production functions, we can understand one of the most famous laws in all economics, the law of diminishing returns:

Under the **law of diminishing returns,** a firm will get less and less extra output when it adds additional

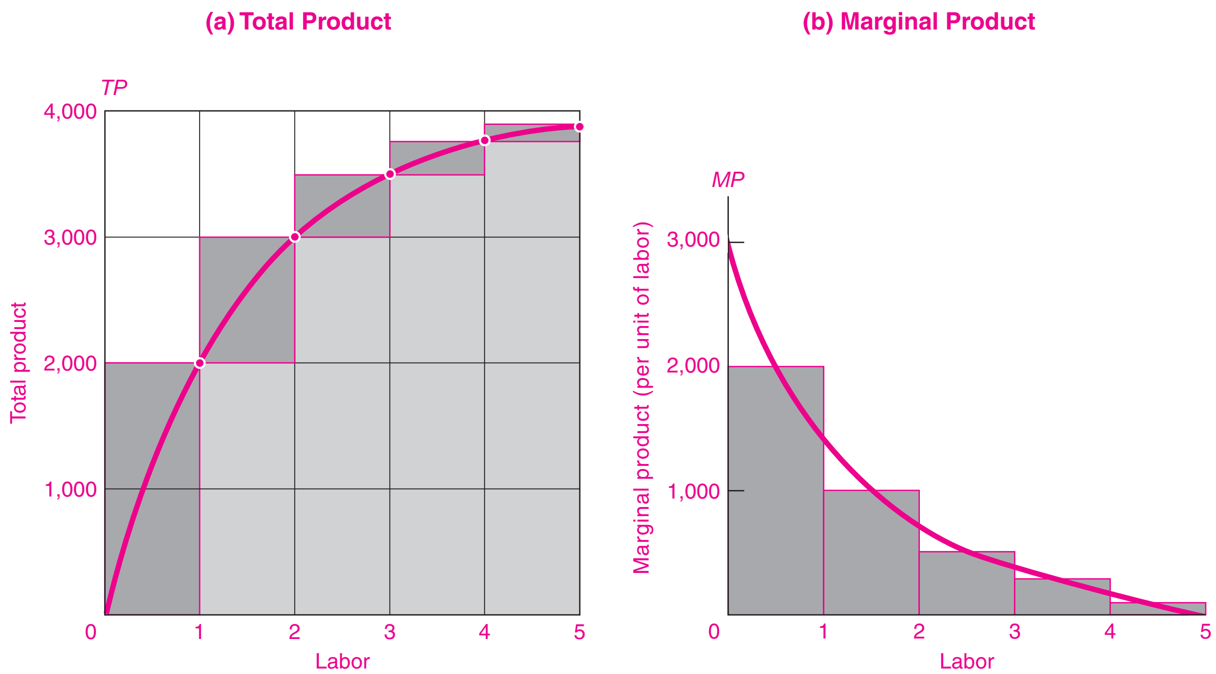


FIGURE 6-1. Marginal Product Is Derived from Total Product

Diagram **(a)** shows the total product curve rising as additional inputs of labor are added, holding other things constant. However, total product rises by smaller and smaller increments as additional units of labor are added (compare the increments of the first and the fifth worker). By smoothing between points, we get the green-colored total product curve.

Diagram **(b)** shows the declining steps of marginal product. Make sure you understand why each dark rectangle in **(b)** is equal to the equivalent dark rectangle in **(a)**. The area in **(b)** under the green-colored marginal product curve (or the sum of the dark rectangles) adds up to the total product in **(a)**.

units of an input while holding other inputs fixed. In other words, the marginal product of each unit of input will decline as the amount of that input increases, holding all other inputs constant.

The law of diminishing returns expresses a very basic relationship. As more of an input such as labor is added to a fixed amount of land, machinery, and other inputs, the labor has less and less of the other factors to work with. The land gets more crowded, the machinery is overworked, and the marginal product of labor declines.

Table 6-1 illustrates the law of diminishing returns. Given fixed land and other inputs, we see that there is zero total output of corn with zero inputs of labor. When we add our first unit of labor to the same fixed amount of land, we observe that 2000 bushels of corn are produced.

In our next stage, with 2 units of labor and fixed land, output goes to 3000 bushels. Hence, the second unit of labor adds only 1000 bushels of additional output. The third unit of labor has an even lower marginal product than does the second, and the fourth unit adds even less. Table 6-1 thus illustrates the law of diminishing returns.

Figure 6-1 also illustrates the law of diminishing returns for labor. Here we see that the marginal product curve in (*b*) declines as labor inputs increase, which is the precise meaning of diminishing returns. In Figure 6-1(*a*), diminishing returns are seen as a concave or dome-shaped total product curve.

(1) Units of labor input	(2) Total product	(3) Marginal product	(4) Average product
0	0		
		2,000	
1	2,000		2,000
		1,000	
2	3,000		1,500
		500	
3	3,500		1,167
		300	
4	3,800		950
		100	
5	3,900		780

TABLE 6-1. Total, Marginal, and Average Product

The table shows the total product that can be produced for different inputs of labor when other inputs (capital, land, etc.) and the state of technical knowledge are unchanged. From total product, we can derive important concepts of marginal and average products.

What is true for labor is also true for any other input. We can interchange land and labor, now holding labor constant and varying land. We can calculate the marginal product of each input (labor, land, machinery, water, fertilizer, etc.), and the marginal product would apply to any output (wheat, corn, steel, soybeans, and so forth). We would find that other inputs also tend to show the law of diminishing returns.

Diminishing Returns in Farm Experiments

The law of diminishing returns is often observed in agriculture. As Farmer Tilly adds more labor, the fields will be more thoroughly seeded and weeded, irrigation ditches will be neater, and scarecrows better oiled. At some point, however, the additional labor becomes less and less productive. The third hoeing of the field or the fourth oiling of the machinery adds little to output. Eventually, output grows very little as more people crowd onto the farm; too many tillers spoil the crop.

Agricultural experiments are one of the most important kinds of technological research. These techniques have been used for over a century to test different seeds, fertilizers, and other combinations of inputs in a successful effort to raise agricultural productivity. Figure 6-2 shows the results of an experiment in which different doses of phosphorus fertilizer were applied on two different plots, holding constant land area, nitrogen fertilizer, labor, and other inputs. Real-world experiments are complicated by "random errors"—in this case, due primarily to differences in soils. You can see that diminishing returns set in quickly after about 100 pounds of phosphorus per acre. Indeed, beyond an input level of around 300 pounds per acre, the marginal product of additional phosphorus fertilizer is negative.

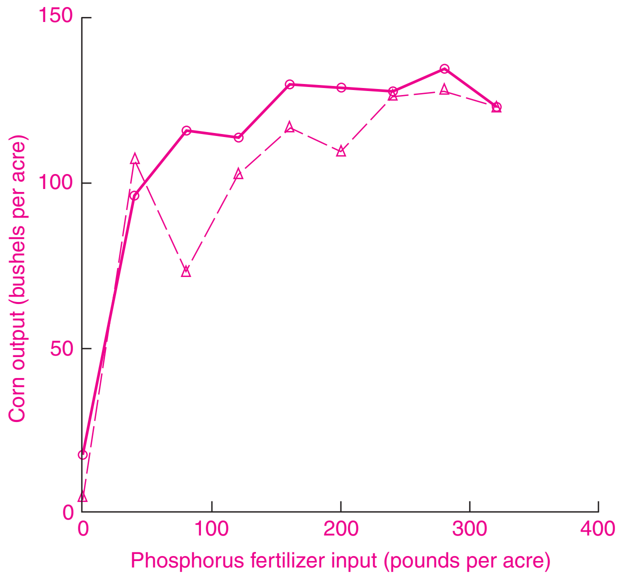


FIGURE 6-2. Diminishing Returns in Corn Production

Agricultural researchers experimented with different doses of phosphorus fertilizer on two different plots to estimate the production function for corn in western Iowa. In conducting the experiment, they were careful to hold constant other things such as nitrogen fertilizer, water, and labor inputs. Because of variations in soils and microclimate, even the most careful scientist cannot prevent some random variation, which accounts for the jagged nature of the lines. If you fit a smooth curve to the data, you will see that the relationship displays diminishing returns for every dose and that marginal product becomes negative for a phosphate input of around 300.

Source: Earl O. Heady, John T. Pesek, and William G. Brown, *Crop Response Surfaces and Economic Optima in Fertilizer Use* (Agricultural Experiment Station, Iowa State College, Ames, Iowa, 1955), table A-15.

Diminishing returns are a key factor in explaining why many countries in Asia are so poor. Living standards in crowded Rwanda or Bangladesh are low because there are so many workers per acre of land and not because farmers are ignorant or fail to respond to economic incentives.

We can also use the example of studying to illustrate the law of diminishing returns. You might find that the first hour of studying economics on a given day is productive—you learn new laws and facts, insights and history. The second hour might find your attention wandering a bit, with less learned. The third hour might show that diminishing returns have set in with a vengeance, and by the next day the third hour is a blank in your memory. Does the law of diminishing returns suggest why the hours devoted to studying should be spread out rather than crammed into the day before exams?

The law of diminishing returns is a widely observed empirical regularity rather than a universal truth like the law of gravity. It has been found in numerous empirical studies, but exceptions have also been uncovered. Moreover, diminishing returns might not hold for all levels of production. The very first inputs of labor might actually show increasing marginal products, since a minimum amount of labor may be needed just to walk to the field and pick up a shovel. Notwithstanding these reservations, diminishing returns will prevail in most situations.

RETURNS TO SCALE

Diminishing returns and marginal products refer to the response of output to an increase of a *single* input when all other inputs are held constant. We saw that increasing labor while holding land constant would increase food output by ever-smaller increments.

But sometimes we are interested in the effect of increasing *all* inputs. For example, what would happen to wheat production if land, labor, water, and other inputs were increased by the same proportion? Or what would happen to the production of tractors if the quantities of labor, computers, robots, steel, and factory space were all doubled? These questions refer to the *returns to scale,* or the effects of scale increases of inputs on the quantity produced. Three important cases should be distinguished:

- **Constant returns to scale** denote a case where a change in all inputs leads to a proportional change in output. For example, if labor, land, capital, and other inputs are doubled, then under constant returns to scale output would also double. Many handicraft industries (such as haircutting in America or handloom operation in a developing country) show constant returns.
- **Increasing returns to scale** (also called **economies of scale**) arise when an increase in all inputs leads to a more-than-proportional increase in the level of output. For example, an engineer planning a small-scale chemical plant will generally find that increasing the inputs of labor, capital, and materials by 10 percent will increase the total output by more than 10 percent. Engineering studies have determined that many manufacturing processes enjoy modestly increasing returns to scale for plants up to the largest size used today.
- **Decreasing returns to scale** occur when a balanced increase of all inputs leads to a less-than-proportional increase in total output. In many processes, scaling up may eventually reach a point beyond which inefficiencies set in. These might arise because the costs of management or control become large. One case has occurred in electricity generation, where firms found that when plants grew too large, risks of plant failure grew too large. Many productive activities involving natural resources, such as growing wine grapes or providing clean drinking water to a city, show decreasing returns to scale.

Production shows increasing, decreasing, or constant returns to scale when a balanced increase in all inputs leads to a more-than-proportional, less-than-proportional, or just-proportional increase in output.

One of the common findings of engineers is that modern mass-production techniques require that factories be a certain minimum size. Chapter 2 explained that as output increases, firms may divide production into smaller steps, taking advantage of specialization and division of labor. In addition, large-scale production allows intensive use of specialized capital equipment, automation, and computerized

Production concept	Definition
Diminishing returns	Declining marginal product of an input, holding all other inputs constant
Returns to scale	Increase in output for balanced increase in all inputs is
Decreasing	. . . less than proportional
Constant	. . . proportional
Increasing	. . . more than proportional

TABLE 6-2. Important Production Concepts

This table shows succinctly the important production concepts.

design and manufacturing to perform simple and repetitive tasks quickly.

Information technologies often display strong economies of scale. A good example is Microsoft's Windows Vista operating system. Developing this program reportedly required $10 billion in research, development, beta-testing, and promotion. Yet the cost of adding Windows Vista to a new computer is very close to zero because doing so simply requires a few seconds of computer time. We will see that strong economies of scale often lead to firms with significant market power and sometimes pose major problems of public policy.

Table 6-2 summarizes the important concepts from this section.

SHORT RUN AND LONG RUN

Production requires not only labor and land but also time. Pipelines cannot be built overnight, and once built they last for decades. Farmers cannot change crops in midseason. It often takes a decade to plan, construct, test, and commission a large power plant. Moreover, once capital equipment has been put in the concrete form of a giant automobile assembly plant, the capital cannot be economically dismantled and moved to another location or transferred to another use.

To account for the role of time in production and costs, we distinguish between two different time periods. We define the **short run** as a period in which firms can adjust production by changing variable factors such as materials and labor but cannot change fixed factors such as capital. The **long run** is a period sufficiently long that all factors including capital can be adjusted.

To understand these concepts more clearly, consider the way the production of steel might respond to changes in demand. Say that Nippon Steel is operating its furnaces at 70 percent of capacity when an unexpected increase in the demand for steel occurs because of the need to rebuild from an earthquake in Japan or California. To adjust to the higher demand for steel, the firm can increase production by increasing worker overtime, hiring more workers, and operating its plants and machinery more intensively. The factors which are increased in the short run are called *variable* factors.

Suppose that the increase in steel demand persisted for an extended period of time, say, several years. Nippon Steel would examine its capital needs and decide that it should increase its productive capacity. More generally, it might examine all its *fixed* factors, those that cannot be changed in the short run because of physical conditions or legal contracts. The period of time over which all inputs, fixed and variable, can be adjusted is called the long run. In the long run, Nippon might add new and more efficient production processes, install a rail link or new computerized control system, or build a plant in Mexico. When all factors can be adjusted, the total amount of steel will be higher and the level of efficiency can increase.

Efficient production requires time as well as conventional inputs like labor. We therefore distinguish between two different time periods in production and cost analysis. The short run is the period of time in which only some inputs, the variable inputs, can be adjusted. In the short run, fixed factors, such as plant and equipment, cannot be fully modified or adjusted. The long run is the period in which all factors employed by the firm, including capital, can be changed.

That Smells So Good!

The production processes of a modern market economy are extraordinarily complex. We can illustrate this with the lowly hamburger.

As Americans spend more time in the workplace and less in the kitchen, their demand for prepared food has risen dramatically. TV dinners have replaced store-bought carrots and peas, while hamburgers bought at McDonald's now number in the billions. The move to processed foods has the undesirable property that the food—after being washed, sorted, sliced, blanched, frozen, thawed, and reheated—often loses most of its flavor. You want a hamburger to smell and taste like a hamburger, not like cooked cardboard.

This is where the "production of tastes and smells" enters. Companies like International Flavors and Fragrances (IFF) synthesize the flavor of potato chips, breakfast cereals, ice cream, cookies, and just about every other kind of processed food, along with the fragrance of many fine perfumes, soaps, and shampoos. If you read most food labels, you will discover that the food contains "natural ingredients" or "artificial ingredients"—such compounds as amyl acetate (banana flavor) or benzaldehyde (almond flavor).

But these unfamiliar chemicals can do amazing things. A food researcher recounts the following experience in the laboratories of IFF:

> [After dipping a paper fragrance-testing filter into each bottle from the lab,] I closed my eyes. Then I inhaled deeply, and one food after another was conjured from the glass bottles. I smelled fresh cherries, black olives, sautéed onions, and shrimp. [The] most remarkable creation took me by surprise. After closing my eyes, I suddenly smelled a grilled hamburger. The aroma was uncanny, almost miraculous. It smelled like someone in the room was flipping burgers on a hot grill. But when I opened my eyes, there was just a narrow strip of white paper.[1]

This story reminds us that "production" in a modern economy is much more than planting potatoes and casting steel. It sometimes involves disassembling things like chickens and potatoes into their tiny constituents, and then reconstituting them along with new synthesized tastes halfway around the world. Such complex production processes can be found in every sector, from pharmaceuticals that change our mood or help our blood flow more smoothly to financial instruments that take apart, repackage, and sell the streams of mortgage payments. And most of the time, we don't even know what exotic substances lie inside the simple (recycled) paper that wraps our $2 hamburger.

[1] Eric Schlosser, *Fast Food Nation* (Perennial Press, New York, 2002), p. 129.

TECHNOLOGICAL CHANGE

Economic history records that total output in the United States has grown more than tenfold over the last century. Part of that gain has come from increased inputs, such as labor and machinery. But much of the increase in output has come from technological change, which improves productivity and raises living standards.

Some examples of technological change are dramatic: wide-body jets that increased the number of passenger-miles per unit of input by almost 50 percent; fiber optics that have lowered cost and improved reliability in telecommunications; and improvements in computer technologies that have increased computational power by more than 1000 times in three decades. Other forms of technological change are more subtle, as is the case when a firm adjusts its production process to reduce waste and increase output.

We distinguish *process innovation,* which occurs when new engineering knowledge improves production techniques for existing products, from *product innovation,* whereby new or improved products are introduced in the marketplace. For example, a process innovation allows firms to produce more output with the same inputs or to produce the same output with fewer inputs. In other words, a process innovation is equivalent to a shift in the production function.

Figure 6-3 illustrates how technological change, in the form of a process innovation, would shift the total product curve. The lower line represents the feasible output, or production function, for a particular industry in the year 1995. Suppose that productivity, or output per unit of input, in this industry is rising at 4 percent per year. If we return to the same industry a decade later, we would likely see that changes in technical and engineering knowledge have led to a 48 percent improvement in output per unit of input $[(1 + .04)^{10} = 1.48]$.

Next, consider product innovations, which involve new and improved products. It is much more difficult to quantify the importance of product innovations, but they may be even more important in raising living standards than process innovations. Many of today's goods and services did not even exist 50 years ago. In producing this textbook, the authors used computer software, microprocessors, Internet

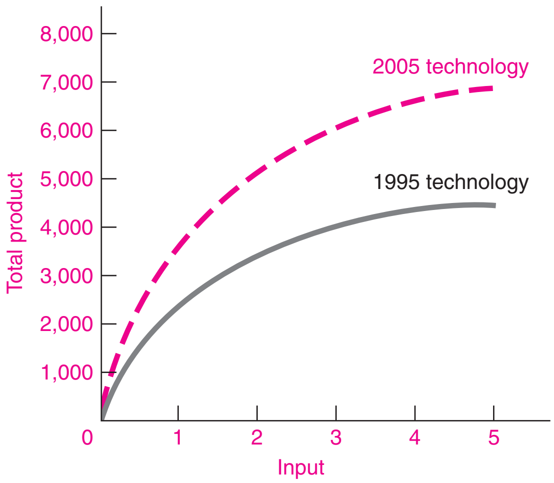


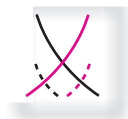
FIGURE 6-3. Technological Change Shifts Production Function Upward

The solid line represents maximum producible output for each level of inputs given the state of technical knowledge in 1995. As a result of improvements in computer technology and management practices, technological change shifts the production function upward, allowing much more output to be produced in 2005 for each level of inputs.

sites, and databases that were not available a decade ago. Medicine, communications, and entertainment are other sectors where product innovations have been critical. The whole arena of the Internet, from e-commerce to e-mail, was not found even in science fiction literature 30 years ago. For fun, and to see this point, try to find any commodity or production process that has not changed since your grandparents were your age!

Figure 6-3 shows the happy case of a technological advance. Is the opposite case—technological regress—possible? The answer is no for a well-functioning market economy. Inferior technologies are unprofitable and tend to be discarded in a market economy, while more productive technologies are introduced because they increase the profits of the innovating firms. To see this, suppose that someone invents an expensive new mousetrap that will never catch a mouse. No profit-oriented firm would produce such a device; and if a poorly managed firm decided to produce it, rational consumers who lived in mouse-infested houses would decline to buy it. Well-functioning markets innovate with better, not inferior, mousetraps.

When there are market failures, however, technological regress might occur. An unregulated company might introduce a socially wasteful process, say, dumping toxic wastes into a stream, because the wasteful process is more profitable. *But the economic advantage of inferior technologies comes only because the social costs of pollution are not included in the firm's calculations of the costs of production.* If pollution costs were included in a firm's decisions, say by pollution taxes, the regressive process would no longer be profitable. In competitive markets, inferior products follow Neanderthals into extinction.



Networks

Many products have little use by themselves and generate value only when they are used in combination with other products. Such products are strongly complementary. An important case is a *network,* where different people are linked together through a particular medium. Types of networks include both those defined by physical linkages, such as telecommunication systems, electricity transmission networks, computer clusters, pipelines, and roads, and the indirect networks that occur when people use compatible systems (such as Windows operating systems) or speak the same language (such as English).

To understand the nature of networks, consider how far you could drive your car without a network of gas stations or how valuable your telephone or e-mail would be if no one else had telephones or computers.

Network markets are special because consumers derive benefits not simply from their own use of a good but also from the number of other consumers who adopt the good. This is known as an *adoption externality.* When I get a phone, everyone else with a phone can now communicate with me. Therefore, my joining this network leads to positive external effects for others. The network externality is the reason why many colleges provide universal e-mail for all their students and faculty—the value of e-mail is much higher when everyone participates. Figure 6-4 on page 115 illustrates how one individual's joining a network has an external benefit to others.

Economists have discovered many important features of network markets. First, network markets are "tippy," meaning that the equilibrium tips toward one or only a