chapter 26

The Long-Run Demand for Labor and Adjustment Costs

I n practice, it takes time—sometimes several years—for firms to increase their capital stocks (by investing in new plant and equipment) or reduce them (by selling off their capital at auction or not replacing it as it depreciates). Given that capital is such an important element of the production process, it is essential that economists understand the firm's behavior both during the period when its capital stock can reasonably be taken as temporarily fixed (the *short run*) and when it is variable (the *long run*). In Chapter 3, we made considerable progress toward achieving this goal when we analyzed the firm's short-run demand for labor (in a variety of different market settings). In this one, we shall fully realize it, by studying the long-run behavior under conditions of perfect competition. As we shall see, the firm's ability to adjust its capital stock can have a profound impact on its behavior because it can now respond to any given impulse—such as an increase in the wage—by responding along two different dimensions.

Sections 26.1–26.3 present the basic theory and offer several applications of the material. As we shall see, the analysis sheds light on several interesting issues such as the consequences of a worldwide prohibition on the use of child labor, trade union activities, and even the behavior of the Luddites.

Section 26.4 then examines how the demand for labor is affected by the presence of *adjustment costs*, which are incurred when workers are hired or fired. For example, if a firm recruits additional workers, then it may face significant outlays when it advertises its vacancies and interviews potential job candidates. Likewise, if it subsequently downsizes, then it may be contractually obliged to provide severance payments to those it lays off. This section shows how the adjustment-cost framework can be applied to help us understand the effects of *job-security provisions*, which often result in employers bearing substantial costs whenever they release workers. These provisions are

endemic in the countries of western Europe, and some observers have argued that they constitute the smoking gun that is responsible for their arthritic—or "Eurosclerotic"—economic performance in general and high levels of unemployment in particular.

Finally, Appendix 26.A presents a more detailed microeconomic analysis of the firm's behavior, and, for completeness, Appendix 26.B provides a mathematical derivation of many of the results obtained in the body of this chapter.

26.1 The Competitive Firm

In this section, the general principles that govern the firm's long-run demands for capital and labor are presented. For simplicity, throughout the discussion, the firm is assumed to be a perfect competitor in both its product and factor markets. It therefore treats its product price, p, the hourly wage, W, and the rental price of capital, R, as given when it formulates its plans. Furthermore, all of the conditions presented in Model 3.2 are again assumed to hold but with one obvious caveat: since we are now focusing on the long run, the firm's capital stock is no longer fixed at the level K_0 but is free to vary.

The manager's goal is to find the particular values of y, K, and L—connoted, respectively, by y^*, K^*, and L^*—that maximize the firm's profits. *Our* goal is to discover these optimal choices and characterize their properties. Essentially there are two alternative (but economically equivalent) methods we can use to accomplish this task: the *input approach* and the *output approach*. In this section, we use the first method because it is simpler and more direct. The output approach is presented in Appendix 26.A. Although it is a little more complicated, it provides much deeper insights into the firm's long-run behavior.

The Input Approach

The input approach is based on the fact that the firm's input choices automatically determine its output level via its production function: $y = F(K, L)$. It follows that its revenues are then $py = F(K, L)$, and that its profits can then be written as:

$$\Pi = p \cdot F(K, L) - WL - RK \tag{26.1}$$

Equation 26.1 implies that the manager's problem just boils down to finding the profit-maximizing input levels, K^* and L^*—hence our designation the input approach. This endeavor is straightforward if he adheres to the one-step-at-a-time principle discussed in Chapter 3. The main result is presented in Major Result 26.1, and the explanation follows.

MAJOR RESULT 26.1

The Long-Run Demand for Capital and Labor

The firm's optimal long-run behavior is described by:

$$p \cdot MP_L = W \qquad (26.2a)$$
$$p \cdot MP_K = R \qquad (26.2b)$$
$$F(K^*, L^*) = y^* \qquad (26.2c)$$

where MP_L and MP_K are, respectively, the marginal products of labor and capital.

One of the main findings obtained in Chapter 3 (see Major Result 3.4) is that, the competitive firm's profit-maximizing demand for labor necessarily satisfies $p \cdot MP_L = W$, for *any* given fixed level of the capital stock K_0. Equation 26.2a says that this is also clearly true if the particular given capital equals its optimal long-run level K^*.

Equation 26.2b has an analogous interpretation, but it governs the firm's optimal choice of capital. It is again derived using the one-step-at-a-time principle, and it ensures there is no scope for the firm to increase its profits by making marginal adjustments in its capital stock. Finally, Equation 26.2c basically reminds us that the firm's optimal output level depends directly on its optimal input levels.

Equations 26.2a–26.2c constitute three equations in three unknowns, and they can be used to determine the firm's profit-maximizing choices of labor, capital, and output. Worked Problem 26.1 shows how this is done in practice.

The Equal-Bang-for-the-Buck Condition. Major Result 26.1 also offers valuable insights into the *general* properties of the firm's long-run behavior. To see how, notice that Equation 26.2a and Equation 26.2b can be rearranged as follows:

$$\boxed{MP_L/W \equiv MP_K/R} \qquad (26.3)$$

This condition properly accounts for the cost and productivity of each input.

ECONOMIC APPLICATION 26.1

Myths about Productivity, Payments, and "Cost Cutting"

The equal-bang-for-the-buck condition, Equation 26.3, provides valuable insights into the business of—well—running a business. For instance, it is common to hear that, in today's high-tech world, firms should hire only the best workers together with the most technologically cutting-edge capital equipment. Yet, as Equation 26.3 makes clear, the validity of this claim depends on carefully comparing the costs and the benefits of each input.

For instance, because of their reliability and speed the latest high-end, H, PCs generate an impressive $10,000 per month in revenues at a cost of only $5,000 per unit. In contrast, medium-range, M, computers generate what appears to be the apparently paltry sum of $2,000 in revenues. However, if the price p^M is less than $1,000 per unit, then it makes sense for the firm to purchase the medium-range machines since, under these conditions, $10,000/$5,000 = 2 < $2,000/p^M$. For instance, if p^M = $500 then five new medium-quality machines generate $10,000 = 5 \times $2,000 in revenues at a cost of only $2,500. This compares with one top-end machine, which admittedly generates $10,000, but does so at a *greater* cost of $5,000.

In a similar spirit, it is sometimes claimed that a firm in financial distress should cut its costs by purchasing only *low-cost* capital and hiring *cheaper* low-skilled workers. However, as shown in Worked Problem 26.2 this is potentially disastrous advice because it may actually *raise* the firm's costs, leading to its bankruptcy. Once again the key is Equation 26.3, which weighs the cost of each input with its contribution to the total output. ■

The hourly wage, W, is the number of dollars required to hire *one more* labor hour, so $1/W$ is the number of labor hours the firm can hire if it spends one more dollar on labor. Given the MP_L equals the additional output that is generated from

Worked Problem 26.1

Investment Subsidies and the Long-Run Demand for Labor

Investment subsidies are often viewed as useful policy instruments for stimulating business activity and fostering job creation in depressed urban neighborhoods. In this worked problem we explore how they affect the firm's demand for labor. (Note: This problem does not require knowledge of calculus; however, it does require that the reader can successfully rearrange equations.)

Problem. Let us return to the case of Betsy's pizza parlor (Example 3.1) and suppose her pizza production function is

$$y = F(K, L) = 4K^{0.25} L^{0.25} \qquad (a)$$

where y is her hourly production of pizzas.

(a) Determine her optimal choices of K, L, and y, given p = $10 per pizza, and the hourly wage and rental price of capital are W = $8 and R = $2, respectively.

(b) How would she respond if the government offered a 50% subsidy on her capital expenditures?

Hint. It can be shown (using calculus) that $MP_L = K^{0.25} L^{-0.75}$ and $MP_K = K^{-0.75} L^{0.25}$.

Solution. In tackling this kind of problem it is best to derive the general solution first and then substitute for the particular values of p, R, and W as required.

With the aid of the hint, Equations 26.2 can be written:

$$K^{0.25} L^{-0.75} = w$$
$$K^{-0.75} L^{0.25} = r \qquad (b)$$

where $w \equiv W/p$ and $r \equiv R/p$ are the real wage and the real rental rate of capital, respectively. Dividing the first equation by the second yields:

$$w/r = K/L \rightarrow K = (w/r) \cdot L \qquad (c)$$

Now use Equation c to substitute for K in Equation a:

$$w = (w/r)^{0.25} L^{0.25} \cdot L^{-0.75} = (w/r)^{0.25}/\sqrt{L}$$

Squaring both sides and rearranging yields the desired solution: $L^* = (1/w^2)\sqrt{w/r}$. Using this result in Equation c gives the optimal capital stock, $K^* = (1/r^2)\sqrt{w/r}$. (Notice that the solutions for capital and labor are symmetrical.) Finally, substituting the solutions for K^* and L^* into the production function (Equation a) yields the firm's optimal output level

$$y^* = 4 \cdot (K^*)^{0.25} \cdot (L^*)^{0.25} = 4/\sqrt{rw}$$

(a) Using these solutions and the facts that $p = \$10$, $W = \$8$, and $R = \$27$

$$
\begin{aligned}
L_0^* &= & (1/0.64)\sqrt{4} & = 3.125 \\
K_0^* &= & (1/0.04)\sqrt{0.25} & = 12.5 \\
y_0^* &= & 4/\sqrt{0.16} & = 10
\end{aligned}
$$

(b) A 50% subsidy on capital expenditures implies that the effective rental rate of capital is only $R(1 - 0.5) = 0.5R$ per machine hour. This, the solutions derived earlier, and the facts that $p = \$10$, $W = \$8$, and $R = \$2$ imply

$$L_1^* = 4.4, \quad K_1^* = 35.3, \quad \text{and} \quad y_1^* = 14.1$$

Notice that the subsidy raises Betsy's optimal employment level, so capital and labor are *gross complements*.

the employment of *one* more labor hour, it follows $(1/W) \times MP_L = MP_L/W$ equals the additional output the firm can produce if the firm spends another dollar on labor. Likewise, MP_K/R equals the additional output it can produce if it spends another dollar on capital. From this vantage point, Equation 26.3 then tells us that the profit-maximizing firm acts in a way that, in terms of its output, gets the same **bang-for-the-buck** from each of its inputs!

The equal-bang-for-the-buck condition, Equation 26.3, is both versatile and powerful. Its versatility stems from the fact that it can readily be extended to encompass any number of different inputs—not just two. Economic Application 26.1

Worked Problem 26.2

Cheap Unskilled vs. Expensive College-Educated Labor

Problem. A firm can hire skilled, college-educated, workers for $40 per hour and unskilled workers for $4. Assume that each college-educated worker produces a constant 80 units of output per hour, and each unskilled worker produces 6 units.

(a) Will the firm hire college-educated or unskilled workers?

(b) If the firm produces 960 units of output per hour, what is the cost saving of making the correct hiring decision?

Solution. (a) The temptation appears to be for the firm to try to keep its costs low, by hiring the apparently cheaper unskilled labor. Yet, the equal-bang-for-the-buck condition tells us this is incorrect. To see why, notice if the firm spends an additional $1 on skilled labor it generates $80/40 = 2$ extra units of output. If, instead, it allocates the dollar toward hiring unskilled labor, then it generates only $6/4 = 1.5$ extra units. It is therefore optimal for the firm to hire skilled, college-educated workers.

(b) In order to produce 960 units of output, the firm must hire either $960/80 = 12$ skilled hours (at a cost of $480), or $960/6 = 160$ unskilled hours (at a cost of $640). Therefore, by making the correct choice, the firm saves $160 per hour.

demonstrates its power by showing how it can help dispel some common misconceptions about the best way to run a business, and Worked Problem 26.2 shows how it can be applied to address practical concerns that confront many businesses in real life.

TAKE-HOME MESSAGE 26.1

- In the long run, the firm can vary all of its factors of production, including its capital stock.
- This additional flexibility can have a profound effect on its demand for labor because it can respond to any given change by adjusting along two or more dimensions.
- The firm's input choices pin down its output level via its production function, $y = F(K, L)$. The input approach exploits this fact by writing the firm's profits as $\Pi = p \cdot F(K, L) - (wL + RK)$, which depend on only the input levels K and L.
- One of the main insights yielded by the input approach is that the firm's optimal capital and employment levels are governed by the equal-bang-for-the-buck condition: $MP_L/W \equiv MP_K/R$. This condition properly accounts for the costs and (marginal) productivities of each of the inputs.

26.2 Properties of the Long-Run Demand for Labor

In the last section, we used the input approach to derive the general properties of the firm's long-run behavior. In this one, we shall build on these findings to make predictions about the nature of the firm's long-run demand for labor, compare its long- and short-run employment responses to a given change in the wage, and apply the material to several real-world settings—including understanding the behavior of the Luddites and the economics of child labor.

The Basic Principles

It is possible to identify three key principles that govern the firm's long-run behavior. All of them, in one way or another, essentially confirm our commonsense ideas about how the firm might be expected to behave given its ability to adjust to two or more inputs. The principles are summarized in Major Result 26.2.

MAJOR RESULT 26.2

The Long-Run Demand for Labor

In the long run, the firm responds to a decrease in the wage, W, by

(1) Scale Effect raising its output level and, as a result, increasing its demand for both labor and capital.

(2) Substitution Effect: Two inputs producing any **given** quantity of output using more labor and less capital.

(3) Substitution Effect: Three or More Inputs producing any given quantity of output using (a) more labor and (b) less of at least one other input. Provided these two conditions are satisfied, it might also use more of some other inputs.

The scale effect refers to the fact that a reduction in the hourly wage rate lowers the firm's (marginal) production costs, which encourages it to increase its output (i.e., scale) and to hire more of every input in the process. In contrast, the substitution effect captures the idea that a reduction in the hourly wage lowers the relative cost of labor vis-à-vis capital. The profit-maximizing firm will attempt to take advantage of this, by switching (i.e., substituting) toward the cheaper input. For example, Betsy might respond to a reduction in the wage by hiring two people to wash the dishes rather than just hiring one and investing in an industrial-strength dishwasher.

Notice that effects 1 and 2 work in tandem, so that the long-run demand for labor is predicted to unambiguously increase as the wage W decreases—that is, the long-run labor-demand curve is downward sloping. However, the reduction in the wage has an ambiguous effect on the firm's demand for capital because the scale and substitution effects are in conflict: the demand for capital tends to increase because of the scale effect but decrease because of the substitution effect. If the outcome of this tug of war is that the firm ultimately demands less capital (i.e., the substitution effect dominates), then capital and labor are termed gross substitutes. Alternatively, if the firm demands more capital (i.e., the scale effect dominates), then labor and capital are called gross complements. In this latter case, a reduction in the wage raises the demand for both labor and capital.

The third result is interesting because it says that if a firm produces a given level of output using at least three inputs—say, capital, labor, and energy—then it could respond to a reduction in the wage by reducing its demand for capital and increasing its demand for both labor and energy. If this were indeed the case, then capital and labor would be called net substitutes, and labor and energy net complements.

The Short- vs. Long-Run Demand for Labor

It is instructive to compare the size of the firm's short- and long-run employment responses to a given change in the wage. Here the key result is provided by the **Le Châtelier-Braun principle**, which establishes that the response to a given change in the wage is greater in the long run than in the short run.[1]

Figure 26.1 explains why this is the case. Given the wage W_0, suppose that the firm's long-run profit-maximizing capital and employment levels are, respectively, K_0^* and L_0^*. The firm's optimal employment level is depicted at point A, which lies on the soon-to-be-constructed long-run labor-demand curve. It is essential that,

FIGURE 26.1   The Long-Run Demand for Labor
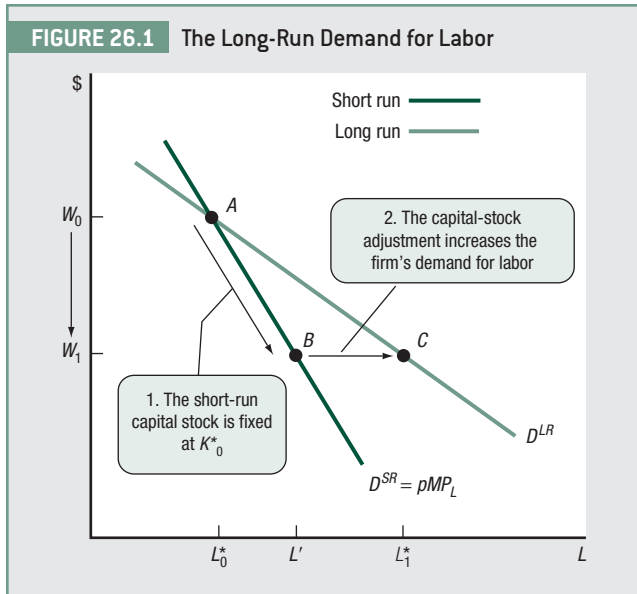
for the moment, the reader completely disregard the other points that lie on the long-run demand curve, D^{LR}, because our goal is to construct it from the basic first principles.

With this goal in mind, starting from point A, suppose that we fix the firm's capital stock at the level K_0^* and we gradually lower the wage to W_1. In this case, because the firm is saddled with the capital stock K_0^*, it must make all of its employment adjustment along its short-run labor-demand curve, $D^{SR} = p \cdot MP_L$. As shown at point B, the firm's short-run demand for labor increases to L'.

In the long-run, the firm responds to the reduction in the wage by adjusting its capital stock and readjusting its employment level. For the reader who has successfully ignored the line D^{LR} depicted in the figure, the basic

question at hand is simple enough: at the new wage W_1 does the firm's long-run demand for labor lie to the left or to the right of point B? If we can show it lies to the right, then we have proven the claim: the long-run labor-demand curve is shallower than the short-run demand curve, which implies the long-run response to a given wage change is greater than the short-run response.

The key to the argument is establishing that, in the long run, the firm optimally adjusts its capital stock in a manner that tends to further increase its demand for labor. Below are the two key principles we will employ to establish this:

- **Gross Complements** If capital and labor are gross complements, then the firm responds to a reduction in the wage by hiring more capital. Moreover, capital and labor are gross complements only if an increase in the capital stock raises the marginal product of labor. (Intuitively, complements go together, so an increase in one input raises the productivity of the other.)

- **Gross Substitutes** If capital and labor are gross substitutes then the firm responds to a reduction in the wage by hiring less capital. Moreover, capital and labor are gross substitutes only if an increase in the capital stock lowers the marginal product of labor. (Intuitively, substitutes go in opposite directions, so an increase in one input lowers the productivity of the other.)

We can make relatively short work of the rest of the argument by using the following suggestive notation to describe the chain of events: let ↑ represent an increase in, ↓ a decrease in, and → implies or leads to. Thus, starting from point B, if capital and labor are gross complements, then the wage cut unleashes the following chain of events:

$$(\downarrow W) \rightarrow (\uparrow K) \rightarrow (\uparrow MP_L) \rightarrow (\uparrow L) \qquad (26.4)$$

> *Tip!*
>
> *Remember, the short-run demand for labor is given by* $W = p \cdot MP_L$. *Hence our goal is to show that as the firm adjusts its capital stock, the* MP_L *increases.*

which takes the firm from point *B* to *C*. Intuitively, the increase in the firm's capital stock raises the marginal productivity of labor and so the firm's demand for labor. Alternatively, in the case of gross substitutes, the chain of events (again starting from point *B*) is

$$(\uparrow W) \rightarrow (\downarrow K) \rightarrow (\uparrow MP_L) \rightarrow (\uparrow L) \tag{26.5}$$

which again takes the firm from point *B* to *C*. This time, it is the reduction in the capital stock that raises the marginal productivity of labor and so the firm's demand for it.

It follows that it is immaterial whether capital and labor are gross complements or gross substitutes. In either case, following the reduction in the wage, the new optimal long-run demand for labor is located at point *C*, which lies on the firm's long-run labor-demand curve, D^{LR}, to the right of point *B*. Connecting points *A* and *C* establishes that the long-run schedule D^{LR} is shallower than the short-run demand schedule D^{SR}. This implies the firm's response to the given wage cut is greater in the long run than in the short run, which confirms the validity of the Le Châtelier-Braun principle.

The Luddites: A Case of Missing the Scale Effect?[2]

The Luddites were bands of workingmen who rioted in the industrial heart of England between 1811 and 1816. The disturbances began in Nottinghamshire, where groups of textile workers in the name of a (possibly) mythical figure called Ned Ludd, or King Ludd, destroyed machinery, to which they attributed their low wages and their high levels of unemployment. In 1812 the riots spread to Yorkshire and Lancashire, where workers wrecked powered cotton looms and wool shearing machines.

Our analysis suggests that the Luddites were spot on in identifying the substitution effect. According to the basic theory, the introduction of low-cost machinery (capital) is predicted to reduce the demand for labor with all else constant. The trouble with the Luddite argument is that it ignores the scale effect. As the price of capital falls, firms are predicted to expand their output levels, which has the offsetting effect of raising the demand for labor.

This observation might help explain the short-lived nature of the Luddite rebellion. Indeed, Tauman and Weiss (1987), and Dowrick and Spencer (1994) show that provided the product market is reasonably competitive, even unions (who value both jobs and wages) often encourage the introduction of new technologies into the workplace.

Child Labor

It is perhaps with some horror that one learns that in 2000 there were some 211 million children, aged between 5 and 14, who were at work worldwide and that 73 million of them were less than 10 years old.[3] It is with perhaps equal

horror that one discovers that during the British industrial revolution, "[E]mploy-ers preferred child labor over adult labor because children were particularly suited to operate the machines."[4] (Which, reading between the lines, means they were more docile and nimble enough to climb into moving industrial machinery in order to maintain and repair it.)

Disturbing as they are, they are simply the facts. It is economic theory that provides insights into the reasons for using child labor in the first place and sheds light on the effectiveness of various policies intended to ameliorate matters.[5] A policy that has gained considerable prominence is the International Programme on the Elimination of Child Labor (IPEC), which seeks to eliminate the worst sorts of child labor abuses.[6]

Analysis.   In order to study the use of child labor, it is necessary that we include it in the firm's production function. Therefore, suppose that the production func-tion takes the form: $y = F(K, L, L_c)$, where K is capital, L is adult labor, and L_C is child labor. A law that restricts the use of children is effectively the same as an increase in the price of child labor. For example, a complete prohibition is eco-nomically equivalent to an infinite price. Like any other factor of production, an increase in its own price will reduce the demand for child labor, raise the demand for inputs that are gross substitutes, and lower the demand for those that are gross complements.

The evidence on the relationship between the use of child and adult labor in production is rather mixed and appears to differ according to gender. For instance, Ray (2000) finds that in Peru an increase in the adult male wage significantly re-duces the labor hours worked by girls; however, in the case of Pakistan he shows there is a strong positive complementarity between the number of hours worked by women and the number worked by girls.[7] The significance of these facts is that if adult and child labor are (gross) complements in production, then an increase in the price of child labor will also reduce the demand for adult labor, which will tend to depress the adult wage. However, exactly the opposite is true if adult and child labor are gross substitutes because the ban will tend to raise the adult wage.

Basu (2000) establishes the possibility of multiple equilibria in this latter case.[8] Intuitively, this means that the labor market will settle down and occupy one of several distinct stable states, and after it does so, left undisturbed, there will be no tendency for any further change. The presence of multiple equilibria raises the distinct possibility that the labor market might become stuck in an undesirable equilibrium, even if other more desirable ones exist that everyone prefers. Fig-ure 26.2 depicts the basic ideas. For simplicity, adults and children are assumed to supply their labor inelastically and to work for a single time period. (The latter as-sumption implies that the wage equals earnings, making it easy to represent them both in the same graph.)

In Basu's model, the family inherently doesn't want to send its children to work but must because of its acute poverty. To capture this idea, suppose that the family

FIGURE 26.2   The Effects of Banning Child Labor
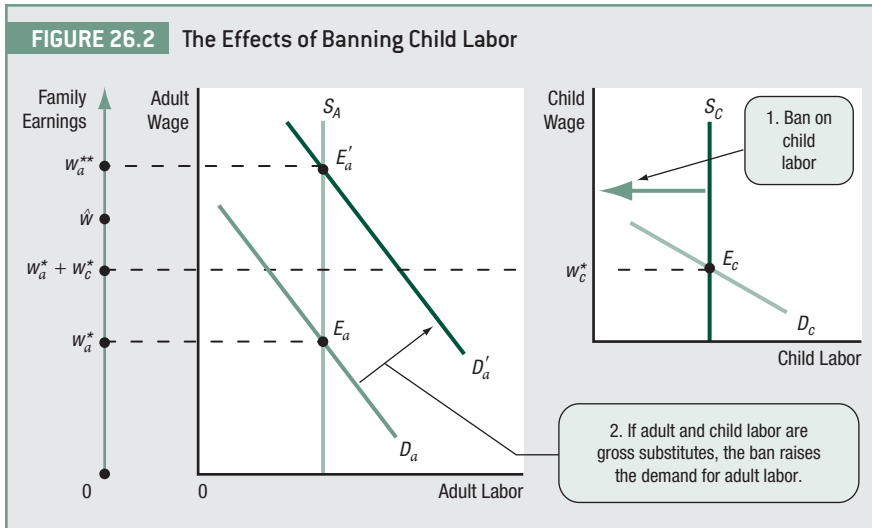
makes its children work only if its total earnings fall below the critical threshold value $\hat{w}$. The initial supply equals demand equilibrium is represented at points E_a and E_c in the figure. The adult wage is w_a^* and the child wage is w_c^*. Notice that $w_a^* + w_c^* < \hat{w}$, which implies the family's earnings are *less* than the critical threshold $\hat{w}$, so it sends its children to work in order to make ends meet.

Now consider the effects of a complete ban on child labor. If adults and labor are (gross) *substitutes*, the ban *raises* the demand for adult labor. This is shown in the figure by the shift in the adult labor-demand schedule from D_a to D_a'. The increase in the demand for adult labor raises the equilibrium adult wage to w_a^{**}, which (in this particular example) exceeds the critical value, $\hat{w}$. Remarkably, in the new equilibrium, the family wouldn't send its children to work even if it were legal to do so. Hence once the labor market reaches point E_a', it would remain there even if the child labor law were subsequently repealed!

TAKE-HOME MESSAGE 26.2

- A reduction in an input price unleashes scale and substitution effects. The scale effect arises because the firm responds to the change by expanding its output level and thus hiring more capital and labor. The substitution effect arises because the reduction in the price of the input encourages the firm to switch its production methods to use more of the relatively cheaper alternative.

- The long-run demand for labor schedule is downward sloping. Following a reduction in the wage, the scale and substitution effects work in harmony and unambiguously increase the firm's demand for labor.

- A reduction in the wage has a theoretically ambiguous effect on the firm's demand for capital. If the scale effect dominates (implying the firm demands more capital), then capital and labor are termed gross complements. If the substitution effect dominates (implying the firm demands less capital) they are termed gross substitutes.
- According to the Le Châtelier-Braun principle, the responsiveness of the demand for labor to a given change in the wage is greater in the long run than in the short run.

26.3 The Hicks-Marshall Laws

Appendix 26.B discusses the important topics of elasticities and percentage changes. One of the principal findings of elasticity is that a small increase in the wage, W, raises the combined earnings of a group of L workers only if it occurs in the inelastic region of the labor-demand curve. This finding is significant because it suggests a union (or any group of workers acting in concert) will tend to push for higher wages if the demand for their labor becomes more inelastic and possibly offer wage concessions if it becomes more elastic.

Tip!
Elasticities are discussed in Appendix 26.B. Some basic familiarity with this material is essential for understanding the following discussion.

Since the own wage labor-demand elasticity can have a potent effect on union behavior, it is important for us to understand the factors that govern its size. Accordingly, this section presents the famous Hicks-Marshall (HM) laws of derived demand.[9] The four laws are presented in Major Result 26.3.

MAJOR RESULT 26.3

The Hicks-Marshall Laws

The demand for labor is inelastic if any of the following conditions hold:

HM1. It is difficult to substitute labor for other factors of production.

HM2. The demand for the industry's product is inelastic.

HM3. The supply of capital (and other factors) is inelastic.

HM4. Labor costs represent a small share of the firm's total costs.

Note: The laws are often remembered by "the inelastics go together."

The HM laws capture the **total change** in the demand for labor that results from an impulse in the wage, after suitable allowances are made for **general equilibrium** adjustments that occur in both the product and the factor markets.[10] This latter statement is less mysterious than it might sound at first. So far, the product price, p, and the rental price of capital goods, R, have simply been treated as givens from the perspective of the individual firm. Yet, these prices are, of course, not

simply conjured up from thin air; they are themselves endogenous and adjust to ensure supply equals demand in their respective markets.

For instance, the product price, p, depends on the demand and supply conditions within the industry as a whole. To give one topical example, the output of crude oil produced by a typical Texan oil well is (literally) a drop in the bucket when compared to total world production. Here, it makes perfect sense to imagine that the producer is a price taker on the world oil market. Yet, as is evident by the market's recent yo-yo behavior, the price of crude depends on both the total demand and the supply of oil. Analogous remarks apply to the market for plant and equipment, where, for example, the price of robotic machinery depends on both their supply and the demand for them.

In order to understand the significance of these observations, assume that capital and labor are gross substitutes. The firm's long-run labor-demand schedule then takes the following form:

$$L = D\,(\,\underset{(-)}{W}\,,\,\underset{(+)}{p}\,,\,\underset{(+)}{R}\,) \tag{26.6}$$

The sign below each of the variables captures the predicted **direct effect** on the firm's long-run demand for labor of a positive impulse in the variable—holding the two other variables constant. If, however, an impulse in any one of these variables affects all of the firms in the industry, then there will also typically be an **indirect effect** on the long-run demand for labor that works through general-equilibrium adjustments in the other two. It follows that in order to correctly evaluate the overall effect of an impulse in the wage, which is our current focus, it is necessary to accommodate these indirect responses. The Hick's-Marshall laws do precisely this.

Below, in the interest of clarity, a step-by-step approach to explain the economic logic that forms the basis for each of the laws is presented. Nevertheless, it is important to bear in mind that, in practice, all of the effects discussed will typically be present simultaneously.

HM1: Capital and Labor Substitutability. Holding constant p and R, the direct impact of a positive impulse in the wage is that it unleashes scale and substitution effects that work together, leading to an unambiguous reduction in the long-run demand for labor (see Equation 26.6). It is convenient to decompose the own wage labor-demand elasticity into these two separate components:

> Elasticity of demand = Scale elasticity + Substitution elasticity

Scale elasticity refers to the percentage change in the demand for labor that results from a 1% impulse in the wage, which operates through the scale effect—likewise for substitution elasticity.

> *Remark.*
> *We focus on pinpointing the conditions under which the demand for labor is inelastic. The conditions under which the demand is elastic are then just the opposite of those presented here.*

Now suppose there are two firms—labeled A and B—that share the same scale elasticity of −0.8%; however, firm A finds it easy to substitute between capital and labor (its substitution elasticity is −1.6%), but firm B finds it impossible to substitute between them (its substitution elasticity is zero). In this case, firm A's labor demand is elastic: a 1% impulse in the wage leads to a 2.4% reduction in its demand for labor. In contrast, firm B's demand is inelastic: the same 1% wage impulses leads to only a 0.8% reduction in its demand. This example confirms HM1: the demand for labor is inelastic if it is difficult to substitute capital for labor.

HM2: The Product Market. The second Hicks-Marshall condition says the demand for labor tends to be inelastic if the demand for the industry's product is inelastic. This is an indirect general equilibrium effect that works through the product market. To see how it works, suppose there is a positive impulse in the wage. Once again, the direct effect of this impulse is that it reduces the firm's demand for labor—see Equation 26.6 and HM1.

However, as each firm in the industry cuts back its output level, total industry output, Y, declines. Figure 26.3 shows the effect is that the industry's supply curve moves leftward (from S_0 to S_1) along the given product demand curve. In turn, this raises the equilibrium product price, p, which then has the blowback effect of raising the demand for labor (see Equation 26.6) and partially offsetting the direct effect of the wage increase! Using our earlier suggestive notation, this indirect sequence of events can be compactly written in the following form:

$$(\uparrow W) \rightarrow (\downarrow y) \rightarrow (\downarrow Y) \rightarrow (\uparrow p) \rightarrow (\uparrow L) \qquad (26.7)$$

Comparing points E' and E'', it can be seen that the magnitude of the offsetting product price increase is greatest if the industry's product-demand curve is completely inelastic. This confirms the result: inelastic product demand and inelastic labor demand go together.

FIGURE 26.3   The Role of the Product Market
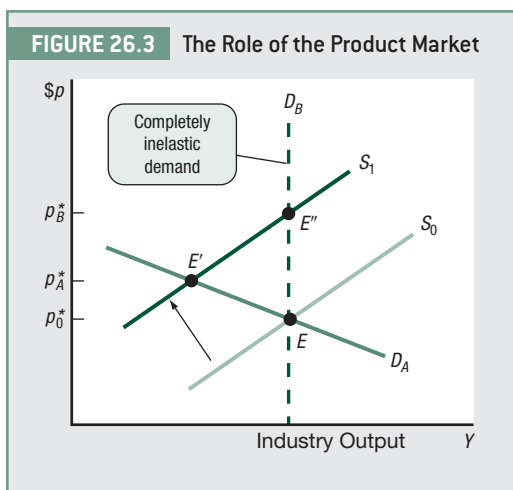
HM3: The Market for Capital Equipment. The third Hicks-Marshall condition is also an indirect general equilibrium effect. This time it operates through the capital equipment market, rather than the product market. To see how it works, once again consider a positive impulse in the wage. The direct effect of the impulse is that each firm in the industry reduces its demand for labor, for the reasons we have already described at length.

Suppose that capital and labor are gross substitutes (the analysis is also valid if they are gross complements—see Worked Problem 26.3). It follows that, all else equal, the increase in the wage indirectly raises each firm's demand for capital and hence the industry's demand for capital as a whole. In turn, this tends to *raise* the

equilibrium machine rental price, R. Since, however, capital and labor are assumed to be gross substitutes, the increase in the equilibrium rental price R tends to raise the demand for labor (see Equation 26.6), which partially offsets the direct effect of the wage increase. The indirect sequence of events can be compactly written as follows:

$$(\uparrow W) \to (\uparrow K) \to (\uparrow R) \to (\uparrow L) \qquad (26.8)$$

The magnitude of this indirect effect depends on the size of the jump in the equilibrium rental price of capital. It is readily verified (see Figure 26.3) that the

---

Worked Problem 26.3

The Hicks-Marshall Conditions

Problem. Assume that capital and labor are gross complements in a particular industry. Suppose the facts are as follows. All else equal, a:

- 2% increase in the wage reduces the demand for labor by 3% and the demand for capital by 2%.
- 1% increase in the price of capital lowers the demand for labor by 0.45%.
- 1% decline in the demand for capital equipment reduces the equilibrium rental price, R, by 2%.

(a) Assuming p and R are given, is the own wage elasticity of labor demand elastic or inelastic?

(b) What is the overall effect on the demand for labor of a 1% increase in the wage, once accommodation is made for adjustments in the equilibrium price of capital? Is the resulting demand for labor elastic or inelastic?

Solution. (a) *Ceteris paribus*, a 2% impulse in the wage reduces the demand for labor by 3%. The own wage elasticity of labor demand is $-3/2 = -1.5$. Hence the demand for labor is elastic because $|-1.5| > 1$.

(b) This part of the question accommodates both the direct effect of the wage increase and the indirect general equilibrium effects that work through the capital equipment market.

The answer to part (a) has already given us the direct effect: it is -1.5% because a 1% increase in the wage reduces the demand for labor by 1.5%. The indirect effect accommodates the following sequence of events:

$$(\uparrow W) \to (\downarrow K) \to (\downarrow R) \to (\uparrow L)$$

The sequence captures the idea that the posited increase in the wage affects the rental price of capital, which, in turn, has a positive blowback effect on the demand for labor.

The data provided in the question allow us to work through this sequence and calculate the size of the indirect effect. Thus a 1% increase in the wage reduces the demand for capital by 1%. In turn, a 1% decrease in the demand for capital reduces the equilibrium rental price by 2%. Finally, a 2% reduction in the rental price must raise the demand for labor by 0.9% (since a 1% reduction in R raises it by 0.45%). In conclusion, the size of the indirect effect is +0.9 because a 1% increase in the wage indirectly raises the demand for labor by 0.9%.

The overall impact of the increase in the wage is found by summing the indirect and direct effects. Consequently, a 1% increase in the wage results in a $(-1.5 + 0.9) = -0.6\%$ change in the demand for labor. It follows, as indicated by HM3, the demand for labor is relatively inelastic once equilibrium adjustments in the capital equipment market are properly accounted for.

increase will be greatest if the capital equipment supply curve is completely in-elastic. Consequently, inelastic labor demand and inelastic capital supply go to-gether, which confirms HM3. Worked Problem 26.3 shows that HM3 is valid even if capital and labor are gross complements, implying an increase in R lowers the demand for labor.

HM4: Labor's Share in Total Costs.   In this case, the Hicks-Marshall conditions predict that the demand for labor will tend to be inelastic if labor's share in total costs is small. The intuition is as follows. If, for example, labor constitutes only 5% of production costs, then a 10% increase in the wage raises total costs by only 0.5%. Under the circumstances, the induced scale effect and resulting decline in employment are small, implying the demand for labor is relatively inelastic. Com-pare this situation with one in which labor is the only factor of production. Here, the same 10% increase in the wage results in a 10% increase in the firm's costs. In turn, this unleashes both a huge scale effect and marked employment decline, implying the demand for labor is relatively elastic.

While this is all well and fine, there is an important caveat: it cannot be too easy to substitute the input in question for another one. For instance, suppose the firm's (homogeneous) workforce is arranged in groups corresponding to the first letter of each worker's surname. In the United States, chances are that the X's would constitute a very small share of the firm's total costs. Yet, it would be a seri-ous mistake for Ms. Xylophone to confidently look at the fourth Hicks-Marshall condition and conclude that (because of her small share in the firm's costs) the demand for her labor must be inelastic and, worse still, then demand a large pay increase from her boss. The reason, of course, is that her employer can, at the drop of a hat, substitute her labor for any Tom, Dick, or, for that matter, Harry.

Elasticities: The Empirical Evidence.   A vast amount of empirical research has been carried out that tackles the difficult problem of estimating labor-demand elasticities. The most comprehensive account of the evidence is presented by Hamermesh (1993). Summarizing the key results as they pertain to homoge-neous groups of workers, Hamermesh remarks,

> We know that the absolute value of the constant-output elasticity of demand for homogeneous labor for a typical firm, and for the aggregate economy in the long run, is above 0 and below 1. Its value is probably bracketed in the interval: [0.15–0.75] with 0.30 being a good "best guess."[11]

By holding output as fixed, the elasticity of labor demand reported above cap-tures the size of the substitution possibilities between labor and other factors of production. The (long-run) own wage elasticity of labor demand measures the response in employment to a 1% change in the wage. Estimates on its value vary considerably from one industry to another. Hamermesh[12] cites evidence from

Carruth and Oswald's (1985) study of UK coal mining employment that suggests in this case it lies in the range $[1.0{-}1.4]$.[13]

Implications for Trade Union Behavior

The Hicks-Marshall conditions can help to illuminate certain aspects of trade union behavior. The linchpin of the argument is grounded in Major Result B.1 (p. A-30). This result is important because it shows that if the own wage elasticity of labor demand is inelastic, then an increase in the wage raises workers' total incomes, $W \cdot L$. Consequently, under these circumstances, the union leadership can vigorously push for a higher wage, even if it means the loss of some jobs. Those union members who keep their jobs are clearly content with an increase in their wages. More subtly, however, since total labor incomes increase, every union member—even accounting for the ones who lose their jobs—potentially benefits from the increase. The upshot is that if the own wage elasticity of labor demand is inelastic then a trade union is predicted to be at its most powerful: it has something to gain and little to lose by forcing an increase in the wage.

More generally, it follows that union power and union truculence will increase following any change that renders the demand for their labor more inelastic. For analogous but opposite reasons, trade unions will become more conciliatory following one that makes their labor demand less inelastic. Hence there are clearly strong grounds for suspecting an intimate link exists between union behavior and the own wage labor-demand elasticity. Yet, because the Hicks-Marshall laws tell us precisely when we should expect the demand for labor to be elastic or inelastic, they can be used to shed light on trade union activities in a variety of different settings.

On this score, HM1 indicates that the demand for union labor is more inelastic the greater the difficulty in substituting capital (or other inputs) for union labor. This observation helps to explain why unions often insist on manning levels, which require that the firm must employ a given number of union workers at each stage of production. Consequently, as the firm expands its scale, it must hire additional union workers. From the firm's perspective, this effectively renders capital and labor complements, which lowers the labor-demand elasticity and enhances the union's power. Similarly, unions often use closed shop agreements to make it difficult (or impossible) for firms to substitute union labor for (possibly cheaper) nonunion workers. Once again, by hindering substitution possibilities the union increases its power and the earnings of its members.

Next, consider the implications of HM2, which says that inelastic product and inelastic labor demand go together. A corollary of this law is that the union can enhance its power if it can successfully engineer a reduction in the price elasticity of the demand for industry's product. Yet, while correct, this is perhaps easier said than done. The demand for the industry's product depends on the behavior of consumers (households, the government, other firms), whereas the union ostensibly has a direct influence only over the wages and working conditions of its

members. However, the union can indirectly affect the demand for the industry's product by influencing the political process.

To give one topical example, consider the United Automobile Workers (UAW). The demand for domestically produced automobiles is highly elastic because consumers can readily purchase foreign-made vehicles. According to HM2, this renders the demand for UAW labor highly elastic and reduces the union's power. To circumvent this problem, the UAW has a strong incentive to lobby for the promulgation of stringent import controls, in the form of tariffs and/or import quotas, in order to make it harder for consumers to switch to purchasing foreign-made vehicles.[14] The reason is that these measures stymie foreign competition, reduce the price elasticity of demand for U.S.-produced automobiles, and (according to HM2) reduce the own wage demand elasticity for UAW labor. In other words, by lobbying for import controls, the union can boost its power and, as a consequence, its members' earnings.

Nevertheless, the political process runs both ways, and legislative changes sometimes weaken unions. The evolution of the U.S. airline industry offers an instructive case in point. Beginning with a major deregulation in 1978, the industry has undergone profound changes over the past 30 years or so. Prior to 1978 many carriers were assigned exclusive rights to fly between certain cities, which gave them de facto monopoly power. Moreover, since driving from Boston to San Francisco is no real substitute for flying, each airline faced an inelastic product demand schedule. As predicted by HM2, this situation would allow, for example, the pilot's union to push for—and obtain—high wages without jeopardizing too many of their jobs.

The 1978 deregulation of the industry, however, eliminated the exclusive rights provision, which essentially led to the evaporation of each carrier's monopoly power. Thus if the ticket price of one carrier was out of line with the others, consumers would "vote with their seats" and choose a cheaper carrier. The resulting increase in competition for passengers is predicted to raise each carrier's product demand elasticity, which, according to HM2, results in a more elastic demand for pilots' labor and weakens their power. This observation helps to explain the recent outcome of negotiations between Delta and its pilots' union, which led to a 32.5% reduction in their earnings.[15]

TAKE-HOME MESSAGE 26.3

- The firm's long-run labor-demand schedule takes the following form: $L = D(W, p, R)$. The direct effect on its behavior of an impulse in any one of these variables is calculated holding the other two fixed.

- If, however, the impulse affects all of the firms in the industry, then there will typically also be an indirect effect on its long-run demand for labor, which works through general-equilibrium adjustments in the other two.

- The Hicks-Marshall laws, which are summarized in Major Result 26.3, accommodate these indirect adjustments.
- A small increase in the wage raises the combined earnings of a group of L workers only if it occurs in the inelastic region of their labor-demand curve. One implication of this observation is that unions will become more truculent and powerful following any change that renders the demand for their labor more inelastic.

26.4 Adjustment Costs

In this section, we will examine how firms optimally alter their employment levels in the presence of adjustment costs. These costs are ubiquitous in practice. For example, when the firm recruits workers it may incur substantial outlays in advertising its vacancies. Likewise, if it sheds labor, then it may be contractually bound to make severance payments. Furthermore employment adjustments often result in a substantial dislocation of the production process, leading to a costly loss of output, as new recruits are trained or as existing employees take over the work once carried out by their former colleagues.[16]
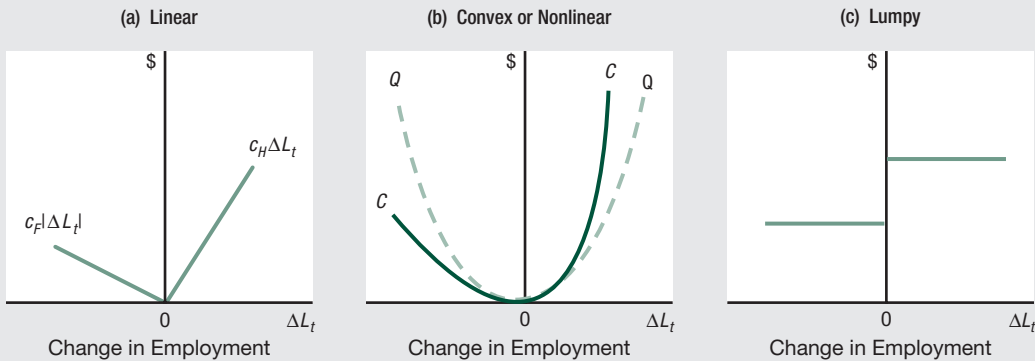
The Nature of Adjustment Costs

Firms must continually make adjustments to their employment levels as business conditions evolve. For instance, during good times, a firm may hire additional workers to meet an increase in the demand for its product; during normal times, it may be obliged to hire workers to replace those who quit or retire; and during bad times, it may be forced to cut its employment level by initiating plant closures or layoffs. Furthermore, in a process called **labor churning**, it is common for firms to simultaneously hire and fire workers. Adjustment costs refer to the costs associated with the loss of existing employees and the recruitment of new ones.

Treadway (1971) was the first to draw the important distinction between **internal** and **external costs** of adjustment.[17] Internal adjustment costs refer to the output losses the firm suffers from the disruption of the accustomed flow of work as it adjusts its employment level. External adjustment costs refer to any pecuniary costs the firm incurs. For example, firms must pay to advertise their vacancies, must incur expenses in training new recruits, and may be contractually bound to pay severance pay to those workers they release. Estimates indicate that external adjustment costs alone are very high, and can amount to as much as one year's payroll for each worker![18]

Three Alternative Adjustment-Cost Structures. In order to characterize the firm's adjustment costs, we must describe the change in its employment level. Accordingly, let $t,\ t + 1,\ t + 2,\ \ldots$ index different time periods (e.g., months); let

FIGURE 26.4   Three Adjustments for Cost Structures

(a) Linear   (b) Convex or Nonlinear   (c) Lumpy

$c_H \Delta L_t$

$c_F |\Delta L_t|$

Change in Employment   Change in Employment   Change in Employment

L_t denote the date-t employment level; and let $\Delta L_t \equiv L_{t+1} - L_t$ denote the net change in employment between t and $t+1$. For simplicity, assume that the firm's adjustment costs depend on only net employment changes. Figure 26.4 depicts three possible adjustment cost structures. We next consider the implications of each of them, in turn, on the firm's optimal behavior.

Linear Adjustment Costs. In the case of linear adjustment costs, the firm incurs the uniform adjustment cost $\$c_H$ if it hires a worker and $\$c_F$ if it fires one. Consequently, it incurs the total adjustment costs $c_H \Delta L_t$ if it hires (on net) $\Delta L_t > 0$ workers, and $c_F |\Delta L_t|$ if it releases $|\Delta L_t| > 0$ workers—remember $|\cdot|$ is the absolute or positive part of the number. This leads to the adjustment cost structure depicted in Figure 26.4a. Notice that the adjustment costs are assumed to be asymmetric, with hiring costs exceeding firing costs.[19]

FIGURE 26.5   Optimal Employment in the Presence of Linear Adjustment Costs

$\$W$

$W_0 + c_H$

W_0

$W_0 - c_F$

Hiring   Firing

MRP_L

L_0^A L_H^* L_0^* L_F^* L_0^B L
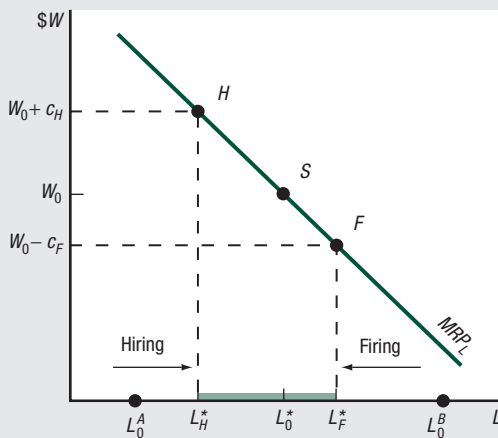
Suppose that the firm possesses a standard downward sloping marginal revenue product schedule MRP_L, and it takes the competitively determined wage, W_0, as given. In the absence of adjustment costs, it would instantly select the profit-maximizing employment level, L_0^*, depicted at point S in Figure 26.5. Now suppose that it faces the adjustment-cost structure depicted in Figure 26.4a.

To begin with, assume that it initially has L_0^A employees. As shown in Figure 26.5, the MRP_L is extremely high, which signals it might want to hire more workers. The cost of hiring an additional worker, however, is

$(W_0 + c_H)$, which includes the hiring cost $\$c_H$. Hence, beginning with L_0^A workers, the optimal level of employment is L_H^*, as this equates the costs and benefits of hiring labor at the margin. Moreover, since there is no point dillydallying, the firm is predicted to immediately adjust its employment level, causing it to jump from L_0^A to L_H^*.

Now suppose that it begins with L_0^B workers, which implies the MRP_L is quite small (see point F in Figure 26.5) and indicates that the firm might want to get rid of some of its employees. However, the presence of the firing adjustment cost implies the marginal cost of labor is $\$(W_0 - c_F)$. The reason is that if it retains an additional worker (i.e., it fires one less), then it must pay the wage $\$W_0$ but avoids the adjustment cost $\$c_F$. The firm's profit-maximizing level of employment is depicted at point F in the figure. Once again, since there is no point dillydallying, the firm's employment level instantly jumps from L_0^B to the new optimal level L_F^*.

Finally, suppose that its initial level of employment lies between L_H^* and L_F^* (and differs from L_0^*). In this case, $W_0 + c_H > MRP_L > W_0 - c_F$, which implies that it is optimal for the firm to leave its employment level unchanged. This contrasts with the zero-adjustment cost case, in which the firm's employment level always instantly jumps to L_0^*, regardless of its initial level. Together, these findings lead us to one of the central insights of the adjustments cost literature.

> **Tip!**
> The fact that we add the hiring cost but subtract the firing cost from the wage is a common source of confusion. Remember, if the firm is downsizing but retains one more worker, then it lays off one less. It must pay the wage but saves the firing cost.

MAJOR RESULT 26.4

Adjustment Costs and Employment Volatility

Employment volatility declines in the presence of adjustment costs.

Nonlinear Adjustment Costs. In the case of linear adjustment costs, it does not matter whether the firm changes its employment level by 1 worker or by 10,000: in each case, the adjustment cost per worker is the same. In practice, however, the disruption to the firm's production process would clearly be much greater in the latter case than in the former, which suggests that adjustment costs might increase rapidly with the size of the adjustment $|\Delta L_t|$. This fact leads to the convex adjustment cost structures shown in Figure 26.4b. Notice that, in each of the cases illustrated, it is not very costly for the firm to make modest employment adjustments (in either direction) but very costly for it to make rapid large-scale changes.

The quadratic case, which is represented by the curve QQ, was, until fairly recently, the bread and butter of dynamic labor-demand models (it is very easy to characterize this case mathematically). The problem is that this formulation necessarily implies that adjustment costs are symmetric, so it is equally costly for the firm to hire or fire, say, 100 workers. This is quite unrealistic. Previously, hiring costs were described as resulting from training costs and from the costs of advertising vacancies, whereas firing costs were described as resulting from the

FIGURE 26.6   The Optimal Adjustment of Labor over Time
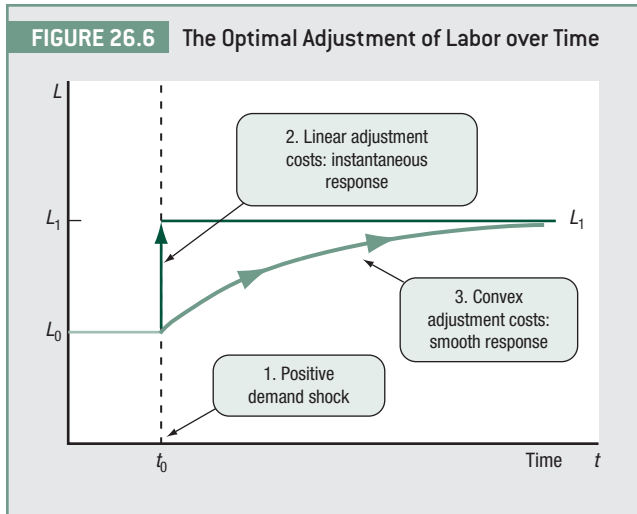
dislocation of production and from severance payments. The fact that these costs are different in kind is enough to make it extremely unlikely that they would happen to equal each other by chance. Accordingly, the solid curve *CC* in the figure illustrates the more realistic case of asymmetric and increasing marginal adjustment costs.

The speed with which the firm optimally adjusts its workforce over time is extremely sensitive to its underlying adjustment cost structure. Figure 26.6 explains the basic principles. The firm is assumed to begin with the initial (optimal) employment level L_0 and is subject to a positive demand shock, at date t_0, which increases it optimal employment level to L_1.

In the case of *linear* adjustment costs, its employment level simply *jumps* to the new value L_1 (if it adjusts at all). The reason is that if it is profitable to hire one more employee then it must be profitable to hire all of the additional $L_1 - L_0$ workers it requires without delay. In the case of convex adjustment costs, however, such rapid adjustments are extremely costly. As a consequence, it is optimal for the firm to smoothly adjust its level of employment until it reaches its new target level L_1. In choosing the speed of its employment response, ΔL_t, the firm weighs the benefits of reaching its final destination, L_1, more rapidly, against the higher adjustment costs that arise if it changes its workforce too quickly.

Lumpy Adjustment Costs. Figure 26.4*c* depicts the case of *lumpy adjustment costs,* which have attracted considerable recent attention. Notice, the costs of adjustment are independent of the number of workers hired or fired during the period. (For example, the cost of running an advertisement in the local paper that says the firm wants to hire 5 new workers is the same as running one that says it wants to hire 25.) Because of their lumpy nature, the optimal level of employment is insensitive to small shocks that affect the MRP_L. However, larger shocks can unleash sudden and dramatic employment changes, by making it worthwhile for the firm to bear the fixed-cost element.

Evidence. Anderson (1993) empirically implements a model with linear adjustment costs that stem from the *experience-rating* feature of the unemployment insurance system.[20] (The experience-rating system implies that the number of workers the firm lays off in one period affects its unemployment insurance tax liabilities in future periods.) She finds adjustment costs dampen each firm's response to shocks and reduce employment variability.

Hamermesh (1989) develops and empirically implements a model, using monthly plant level data, that incorporates lumpy adjustment costs. His results indicate that the assumption of fixed adjustment costs does a better job at explaining the data than the convex adjustment cost case depicted in Figure 26.4*b*. Hamermesh (1992) extends his earlier analysis to encompass both lumpy and variable adjustment costs. He estimates the model using two different data sets: one includes observations on production workers employed in seven large plants that were owned and run by a large manufacturing company; the other, observations on airline mechanics employed at seven trunk airlines. He finds that the lumpy adjustment cost model appears to better characterize the manufacturing plant data but the combined variable lumpy adjustment cost model is better for airline mechanics.

Job-Security Provisions and Eurosclerosis

In many countries—including the United States but especially those in western Europe—employers cannot respond to adverse business conditions by simply shedding labor at will. Instead, their discretionary powers are often limited by legal restrictions and other contractual obligations that collectively fall under the general rubric of **job security provisions**. These provisions often impose serious costs on firms whenever they attempt to release some of their employees.[21]

Examples of these sorts of provisions are not hard to find. In the United States, the 1988 Worker Adjustment and Retraining Notification Act (WARN) requires that firms provide workers with a 60-day advanced warning in the event of substantial layoffs or plant closings. Similarly, firms are often contractually bound to make severance payments to those workers they lay off. Furthermore, in many European economies, workers are often protected against unfair dismissal, and employers must often demonstrate good cause whenever they attempt to fire some of their workers.

To the pundit and policy maker alike the ability of job security provisions to reduce excessive employment volatility and lower the unemployment rate is a self-evident truth. Yet, to economists, who are trained to think carefully about these issues, there is much more to the matter than first meets the eye. The reason is that the effects of job security provisions are often discussed within the context of a simple static environment, in which a group of workers already have jobs and the policy goal is to ensure they keep them. In this setting, promulgating legislation that raises the costs of firing workers does reduce employment volatility, because firms are then inclined to retain their employees during bad economic times. The trouble with this simple story, however, is that we live in a dynamic general equilibrium world, which means that we can't simply ignore how workers got their jobs in the first place! From this general perspective, it is no longer obvious that job security provisions actually work as they are intended.

The key insight is that an increase in the **firing cost**, c_F, can discourage firms from hiring workers in the first place (an increase in the hiring cost, c_H, clearly does so). This is because firms are forward looking: they recognize that if they hire workers in good times, then they may subsequently be forced to release them in bad times and incur a firing cost of c_F per worker. Hence if c_F is sufficiently high, employers may be reluctant to hire additional workers, which could raise the over-all unemployment rate. In fact, in an interesting study, Bertola (1992) examines this issue in more detail. One of his principal findings is that an increase in firing costs, through the enactment of job security provisions, can indeed lower the av-erage level of employment. Worked Problem 26.4 more clearly demonstrates the mechanism in action.

Eurosclerosis. Job security provisions are endemic in the countries of western Europe, and many observers view them as the smoking gun responsible for re-cent lackluster economic performance and high unemployment rates. Thus

Worked Problem 26.4

The Possible Adverse Effects of Job Security Provisions

Problem. It is summer, and a farm intends to tempo-rarily hire some laborers to harvest its strawberry crop. The strawberry collection production technol-ogy is $y = 4800\sqrt{L}$, where y is the quantity collected (in pounds) and L the level of employment. Each pounds sells on a competitive market for $p = \$1$.

(a) How many workers does the farm employ if, over the period, the competitive wage is $W = \$600$ and the firing cost is zero?

(b) Now suppose the government enacts job se-curity provisions that result in the farm bearing a firing cost of $c_F = \$200$ per worker when it releases laborers at the end of the harvesting season. Assum-ing the same wage of $W = \$600$, how many workers does it now employ?

Hint. It can be shown that the marginal revenue prod-uct of labor is $MRP_L = 2400/\sqrt{L}$.

Solution. (a) The farm's profit-maximizing employ-ment level is governed by the condition $MC_L = MRP_L$. Using the hint and the fact that $W = 600 = MC_L$, we

have $600 = 2400/\sqrt{L}$. This implies the optimal level of employment is $L_0^* = 16$ workers.

(b) If the farm hires an additional worker, then it must pay the wage $W = \$600$ and anticipates incur-ring the subsequent separation cost of $200. Hence the marginal cost of labor is $MC_L = W + C_H = \$800$. Using the hint, this implies the farm's optimal level of employment is $800 = 2400/\sqrt{L}$, which yields $L_1^* = 9$ workers. By making firing more costly, the farm hires fewer workers in the first place!

Remark. Some readers may be stunned by the fact that we now **add** the firing cost to the wage! The ex-planation is that, in this problem, the farm currently has zero employees, so the marginal cost of labor is the wage plus the soon-to-be-incurred firing cost. Previously, we investigated the case in which a firm was cutting an already established workforce. In this setting, if it retains one more worker, then it neces-sarily fires one less: it pays the wage W, but avoids the firing cost c_F.

Generous unemployment benefits, restrictions on hiring and firing, . . . are thought to have led to rigid "Eurosclerotic"—as dubbed by Giersch (1985)—economies, which could not cope with the big shocks of the 1970s.[22]

Figure 26.7 depicts the unemployment rates for selected European countries and for the United States between 1960 and 1999.[23] Notice the oil shocks, which occurred during the 1970s, mark the turning point in the countries' relative economic fortunes. Prior to this point the average unemployment rate was considerably lower in Europe than in the United States, but, almost without exception, we have seen a reversal of this pattern since the 1970s.[24]

In the 1960s, with a robust labor market, hiring too many workers was a mistake that would hurt a firm for at most a few months. This situation contrasts with the lean European labor markers of the 1970s and 1980s, during which time the punitive nature of firing costs made the prospect of hiring workers a very risky proposition indeed.

Bertola (1990) has formulated an interesting model of labor demand in the presence of adjustment costs. His main finding is that firing costs can rationalize the dynamic employment behavior witnessed in European economies during the 1970s and 1980s. In a similar spirit, Lazear (1990) adduces evidence for 22 countries over a period of about 30 years, indicating the negative employment consequences of severance pay provisions and advanced notice legislation. How quickly do employment levels respond to shocks?

According to the basic adjustment cost model, the rate of employment adjustment varies with the stringency of job security provisions. Burgess, Knetter, and Michelacci (2000) provide a disaggregated analysis of the issues by looking across countries and across certain industries within each of these countries. Their findings indicate (i) there are dramatic differences in adjustment speeds among industries and (ii) the speed of adjustment is (negatively) related to the extent of job security provisions in the economies in question.[25]

Despite the compelling arguments that job security provisions can have the unintended consequences of reducing employment and stymieing economic growth, introducing the necessary economic reforms is no simple political matter.

FIGURE 26.7 Unemployment Rates: 1960–2000
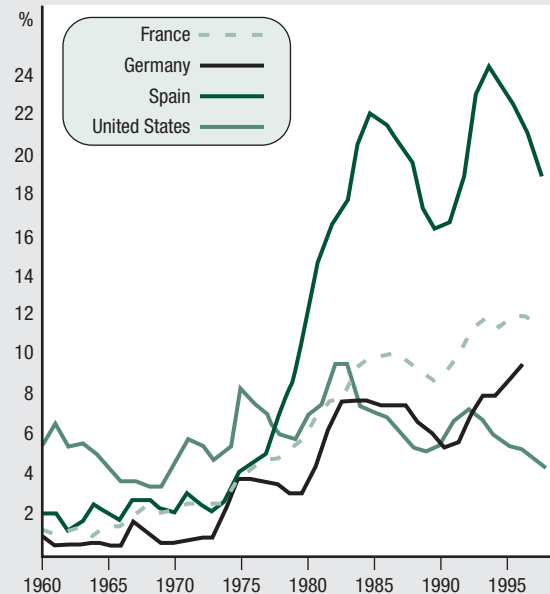
TAKE-HOME MESSAGE **26.4**

- Adjustment costs are incurred whenever a firm hires or releases workers.
- Internal adjustment costs refer to the loss in output that results from dislocations to the normal workflow as the firm changes its employment level. External adjustment costs refer to pecuniary costs that are incurred in the adjustment process.
- One of the central findings of the adjustment cost literature is that these costs tend to reduce employment volatility.
- The way a firm responds to a given shock that affects its optimal employment level depends on its adjustment cost structure. In the case of convex adjustment costs, firms are predicted to smoothly adjust their employment levels over time; in the case of linear or lumpy adjustment costs, firms may respond to small shocks with rapid and large-scale employment changes.
- Job security provisions are an important category of adjustment costs. They refer to legal restrictions and contractual agreements that make it difficult for employers to release workers. Somewhat paradoxically, they can discourage employers from hiring workers and increase the unemployment rate.

SUMMARY

- The long-run demands for labor and capital are governed by the equal-bang-for-the-buck condition:

$$MP_L/W = MP_K/R$$

- where MP/input price equals the extra output the firm can produce if it spends another dollar on that input. Intuitively, this expression properly weighs both the price and productivity of the input.

- In the long run, the firm responds to an increase in the price of an input by cutting back its output level, thus demanding less capital and less labor (the scale effect), and by switching its production technique to take advantage of the relatively cheaper input (the substitution effect).

- The long-run demand for labor schedule is downward sloping. Following a decline in the wage, both the substitution and scale effects work in harmony, which results in an unambiguous increase in the firm's demand for labor.

- According to the *Le Châtelier-Braun* principle, the response to a given wage change is greater in the long run than in the short run.

- The *Hicks-Marshall* conditions offer predictions about the magnitude of the labor-demand elasticity. They assert (*ceteris paribus*) labor demand is most inelastic if any of the following conditions hold:

- **HM1** It is difficult to substitute labor for other factors of production.
- **HM2** Product demand is inelastic.
- **HM3** The supply of capital (and other factors) is inelastic.
- **HM4** The cost of labor represents a small share of the firm's total costs.

- Union power is predicted to increase following any change that renders the demand for their labor more inelastic.

- Firms face an assortment of adjustment costs as they vary their employment levels. Internal adjustment costs refer to those that arise because of disruptions in the regular flow of production. External adjustment costs are pecuniary costs that arise independently of the production process.

- If a firm's adjustment costs are convex, then it is predicted to smoothly alter its employ-ment levels in response to demand shocks. If its adjustment costs are linear or lumpy, then it may respond to shocks with substantial and sudden employment shifts.

- Job security provisions are an important class of adjustment costs. For instance, WARN requires that (many) employers provide workers with 60 days advance notice of mass layoffs. In many European economies, workers are often protected against unfair dismissal, and employers often have to demonstrate good cause for firing a worker.

- Job security provisions can actually reduce the employment levels. Some observers regard them as the primary culprit responsible for the high rates of unemployment seen in Europe since the 1970s.

KEY CONCEPTS

the short vs. the long run

the input vs. output approach

equal-bang-for-the-buck condition

scale and substitution effects

gross vs. net substitutes and complements

Le Châtelier-Braun principle

Luddites

child labor

multiple equilibria

Hicks-Marshall laws

general equilibrium

indirect vs. direct effect

adjustment costs:

1. internal vs. external
2. symmetric vs. asymmetric
3. linear vs. convex vs. lumpy

labor churning

job security provisions

Eurosclerosis

isoquant

marginal rate of technical substitution (MRTS)

isocost

cost minimization

output expansion path

normal input

perfect complements and substitutes

REVIEW QUESTIONS

R1. What is the primary distinction between the short and the long run as it pertains to the firm's demand for labor?

R2. What is the equal-bang-for-the-buck condition?

R3. Explain what is meant by the substitution and scale effects.

R4. How does the long-run demand for labor schedule differ from the short-run demand schedule?

R5. What are the Hicks-Marshall conditions?

R6. What are meant by linear adjustment costs?

R7. Give some examples of job security provisions.

R8. What is an isoquant, and what is meant by the marginal rate of technical substitution?

R9. What is an isocost line?

PROBLEMS

P1. The rapid decline in the price of computers increased the demand for skilled labor workers and reduced the demand for unskilled workers. What does this say about the relationship between capital (i.e., computers) for these two worker groups?

P2. Let W and MP_L denote the wage and the marginal product of labor. Likewise, let R and MP_K denote the corresponding price, and marginal product of capital. What condition determines the optimal choices of these inputs?

P3. College-trained workers generate annual revenues of $100K, per person at a particular firm. They are expensive though: their going wage is $75K per annum. Because of this, the firm is contemplating recruiting much cheaper high-school graduates. Their staring wage is only $16K per worker, but they each generate revenues of $20K. Should the firm hire college-trained workers or high-school graduates?

P4. It is common for U.S. trade unions to be interested in the internal politics of foreign countries, such as in their working conditions and their environmental laws. What economic motive could explain why, for example, the United Steelworkers (USW) might lobby Congress to pass legislation that limits imports from less environmentally friendly countries?

P5. A firm faces linear adjustment costs. The facts are as follows: $W = \$40$, $c_H = \$10$, and $MRP_L = 1100 - 2L$. How does the firm respond if, initially, $L_0 = 475$? What about if $L_0 = 524$?

P6. It is the holiday season, and a firm intends to temporarily open a plant to produce red plastic reindeer noses for the next six weeks (the reader is invited to think of his or her own examples). The nasal production technology is $y = 2400 \sqrt{L}$, where y is the quantity of noses and L the level of employment. Each nose sells on a competitive market for p = $4.

(a) How many workers does the firm employ if, over the period, the competitive wage is $W = \$600$ and both the firing and hiring costs are zero?

(b) Now suppose the government enacts job security provisions that result in the firm bearing a firing cost of $c_F = \$100$ per worker when it closes the plant at the end of the season. Assuming the wage of $W = \$300$, how many workers does it now employ?

Hint. It can be shown that the marginal revenue product of labor is $MRP_L = 4800/\sqrt{L}$.

NOTES

1. The Le Châtelier-Braun principle was first observed in chemical systems. It was introduced into economics by Paul Samuelson in 1949. See en.wikipedia.org/wiki/Le_Chatelier's_principle.

2. Sources: Tauman, Yair, and Yoram Weiss (1987) pp. 477–501; Dowrick, Steve, and Barbara J. Spencer (1994) pp. 316–344.

3. International Labor Organization Report (2002).

4. Tuttle (1999).

5. Basu (1999) pp. 1083–1119.

6. Other policies include Harkin's Bill (1997), which seeks to discourage the use of child labor by banning the import of goods that have been produced using it into the United States.

7. Ray (2000) pp. 3–19.

8. Basu (2000) pp. C50–61.

9. The conditions are named after the renowned British economists John Hicks and Alfred Marshall.

10. A more rigorous account of the laws is contained in Hamermesh (1993).

11. Hamermesh (1993) p. 135.

12. Hamermesh (1993) p. 96.

13. Carruth, and Oswald (1985) pp. 1003–1020.

14. For example, in 2000 the UAW lobbied vigorously against the government's desire to formalize trade relations with China.

15. The Associated Press, Nov. 12, 2004.

16. The seminal paper is by Oi (1962) pp. 538–555. Oi stressed that labor should properly be viewed as a *quasi-fixed factor of production*, in recognition of the fact that employment can be adjusted in the short run but doing so is costly. Nickell (1986); and Hamermesh (1993) provide comprehensive surveys of the dynamic of labor-demand literature.

17. Treadway (1971) pp. 845–855.

18. See Hamermesh (1993) p. 208; and the references cited therein. Barron, Bishop, and Dunkelberg (1985) pp. 43–52, estimate that employers spend over 40 hours training new recruits during the first month of employment; Hamermesh (1993); Hamermesh (1996) pp. 1264–1292, provide a valuable survey of the literature.

19. The evidence indicates that this is the case in practice. See Hamermesh (1996) p. 1268.

20. Anderson (1993) pp. 1015–1042.

21. Bertola (1999) pp. 2985–3024, offers a recent survey of the theoretical implications of job security provisions. See also Bertola (1992) pp. 389–411.

22. Bentolila and Bertola (1990) pp. 381–402.

23. See Van Der Horst (2003).

24. Siebert (1997) pp. 37–54; Nickell (1997) pp. 55–74, provide highly accessible surveys concerning the causes of European unemployment.

25. Burgess, Knetter, and Michelacci (2000) pp. 419–435.

26. The *it can be shown* part of the claim is not difficult to verify. Let Δy, ΔK, and ΔL denote small changes in output, capital, and labor. Using the definitions of the MP_L and the MP_K, we have $\Delta y = MP_K \cdot \Delta K + MP_L \cdot \Delta L$. The trick is noting that along a given isoquant, $\Delta y = 0$, which implies $0 = MP_K \cdot \Delta K + MP_L \cdot \Delta L$. Rearrangement for $\Delta K / \Delta L$ yields the desired result.

27. In other words, with only two inputs, capital and labor must be net substitutes (i.e., holding fixed the firm's output level, a reduction in the wage lowers the demand for capital and raises the demand for labor). This is not necessarily the case in more complicated settings, involving three or more inputs.

28. These *L*-shaped isoquants are characteristic of a Leontief production technology, which was named after its originator Wassily Leontief (winner of the 1973 Nobel Prize in economics).

29. Although suggestive and helpful for building intuition, the example suffers from an unfortunate fatal weakness. While it is true that every bus needs a driver, buses are manufactured in an assortment of different sizes. Hence the bus company can increase its carrying capacity by substituting toward larger buses (i.e., by substituting toward capital), even if it holds the number of bus drivers constant.

REFERENCES

Barron, John M., John H. Bishop, and William C. Dunkelberg. (1985). "Employer Search: The Interviewing and Hiring of New Employees." *Review of Economics and Statistics* 67(1): 43–52.

Basu, Kaushik. (1999). "Child Labor: Cause, Consequence, and Cure, with Remarks on International Labor Standard." *Journal of Economic Literature* 37(1): 1083–1119.

———. (2000). "The Intriguing Relation between Adult Minimum Wage and Child Labour." *Economic Journal* 110(1): C50–61.

Bentolila, Samuel, and Giuseppe Bertola. (1990). "Firing Costs and Labour Demand: How Bad Is Eurosclerosis?" *Review of Economic Studies* 57(1): 381–402.

Bertola, Giuseppe. (1999). "Microeconomic Perspectives on Aggregate Labor Markets." In Ashenfelter, Orley, and David Card, eds. *Handbook of Labor Economics*. Vol. 3C. Amsterdam: Elsevier, pp. 2985–3024.

Bresnahan, Timothy F., and Valerie A. Ramey. (1994). "Output Fluctuations at the Plant Level." *Quarterly Journal of Economics* 109(1): 593–624.

Dowrick, Steve, and Barbara J. Spencer. (1994). "Union Attitudes to Labor-Saving Innovation: When Are Unions Luddites?" *Journal of Labor Economics* 12(1): 316–344.

Hamermesh, Daniel S. (1992). "A General Model of Dynamic Labor Demand." *Review of Economics and Statistics* 74(1): 733–737.

———. (1993). *Labor Demand*. Princeton, NJ: Princeton University Press.

Hicks, John, R. (1966). *The Theory of Wages*. 2nd ed. New York: St. Martin's Press.

Oi, Walter Y. (1962). "Labor as a Quasi-Fixed Factor." *Journal of Political Economy* 70(1): 538–555.

Treadway, Arthur. (1971). "The Rational, Multivariate, Flexible Accelerator." *Econometrica* 39(1): 845–855.

Tuttle, Carolyn. (1999). *Hard at Work in Factories and Mines: The Economies of Child Labor During the British Industrial Revolution*. Boulder and Oxford: Westview Press.

Appendix 26.A The Output Approach

In this chapter, and in Section 26.1 in particular, we characterized the firm's optimal behavior using what we termed the input approach. This approach is both direct and relatively simple. In this appendix, we shall characterize its long-run behavior using the alternative output approach. Although it is a little more complicated, it does provide considerably deeper insights into the economic forces that drive the firm's behavior. But what is the output approach?

The Basic Principle. The starting point for the output approach is to pick any target level of output, $y > 0$, and to (temporarily) treat it as exogenously given. It follows that, given this constraint, if the manager wants to maximize his employer's profits, Π, then he must choose K and L to minimize the costs of producing y, since the firm's revenues are also (temporarily) fixed at $p \cdot y$. If we let $C(y)$ denote the minimum cost of producing y, then the firm's profits can written:

$$\Pi(y) = p \cdot y - C(y) \qquad (26.9)$$

where the notation $\Pi(y)$ shows the dependence of Π on the proposed target level of output, y. It follows from Equation 26.9 that the manager's problem now just boils down to discovering the particular level of output, y^*, that maximizes the firm's profits—hence the designation: "the output approach." In summary, the output approach requires that we carry out the following two steps:

- **Step 1: Cost Minimization.** Determine the input choices that minimize the total costs, $C(y)$, of producing the target level of output y.
- **Step 2: Profit Maximization.** Find the profit-maximizing level of output, y^*.

Step 1: Cost Minimization

The manager's goal is to minimize the total costs of producing the given target level of output y. It is clear that in order for him to do this, he must know (a) the particular input combinations, (K, L) that he can call on to attain the target, y, and (b) the costs associated with each of them. These two observations set the stage for the present discussion.

Production Isoquants. In Chapter 3 (see, in particular, the discussion on page 63) fleeting reference was made to the concept of a **production isoquant**. *Iso* means "equal," and *quant* is short for "quantity." Therefore, an isoquant is the locus of input combinations that result in the production of an equal quantity of output.

Figure 3.1 used a 3D setting to depict the isoquant corresponding to the particular output level y_0. Figure 26.8 depicts the same isoquant y_0—corresponding

FIGURE 26.8   An Isoquant Map
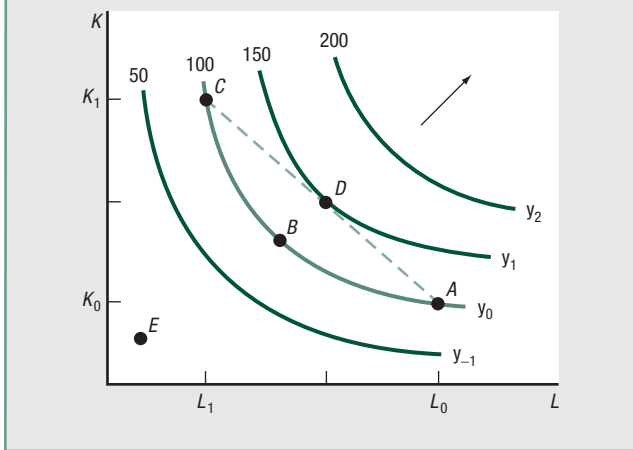
to say 100 pizzas—using a simpler 2D framework. In essence, the figure is derived from Figure 3.1 by pulling out the curve shown there and pasting it into Figure 26.8. Also shown is an assortment of other isoquants that correspond to the different output levels $y_{-1} = 50$, $y_1 = 150$ and $y_2 = 200$ pizzas. Figure 26.8 is called an *isoquant map*, and it compactly summarizes the essential features of the firm's production technology.

Isoquant maps possess several properties that stem from the fact that they are derived from an underlying 3D function like the one shown in Figure 3.1. If in doubt, go back to that figure, and it should be easy to see why the following properties must hold:

- Isoquants do not cross. (How could they? They are derived from different horizontal slices of Figure 3.1.)

- The level of output, y, increases in the direction of the arrow pointing in the NE direction. (Capital and labor are productive inputs, so the more the merrier.)

- An isoquant passes through each and every point shown in Figure 26.8. (Only a handful are ever depicted in the interests of clarity, so, for example, the one passing through point E has been suppressed.)

- Isoquants are negatively sloped and convex to the origin. (The negative slope is apparent, and the convexity property reflects the fact that any chord connecting two points, such as A and C, on a given isoquant lies above the curve.)

The defining feature of a given isoquant is that it represents the locus of inputs that can be used to produce a particular quantity of output. As shown, the firm can produce $y_0 = 100$ pizzas using the input combinations $A(L_0, K_0)$, or $C(L_1, K_1)$, or any other input combination that lies on the curve y_0. In fact, the negative slope of each isoquant reflects the firm's ability to substitute between capital and labor. Hence, starting from point $A = (L_0, K_0)$, if the firm reduces its employment level to L_1, it can still produce y_0 units of output provided it compensates for the decline by increasing its capital stock to K_1 (as shown at point C).

Isoquants are convex because, from the perspective of production, averages are better than extremes. Thus, while the firm can produce y_0 units of output if it uses either the input combinations $A(K_0, L_0)$ or $C(K_1, L_1)$, it can produce the greater level of output $y_1 > y_0$ if it uses the "average"

Tip!

It is common in economics to use the subscripts . . . –1, 0, 1, 2, . . . to label different curves. The convention is that a higher label corresponds to a larger value of the variable, so $y_1 > y_0 > y_{-1}$.

input combination shown at point D. The convexity property also means that the slope of a given isoquant changes as we move along it. With reference to the isoquant y_0, the slope is steep (in absolute value) at point C and relatively flat at point A. Intuitively, this reflects the relative scarcities of capital and labor. For example, at point C capital is abundant but labor is scarce. Therefore, if the firm increases its labor by a unit, then it can shed lots of capital while maintaining its output level constant at y_0 (which is just another way of saying the slope is steep). Similar, but opposite, remarks apply for point A, where capital is the relatively scarce factor.

An isoquant's slope encodes valuable economic information about the firm's production technology. In fact, the slope is so important that economists have a special name for it: the absolute value (i.e., positive part) of the slope is called the **marginal rate of technical substitution of capital for labor**—MRTS for short. The MRTS is measured in terms of *machine hours per labor hour*. For example, suppose that at point A, **MRTS** $= 20$ (machine hours per labor hour). This indicates that if the firm hires one less labor hour, then it must increase its employment of capital by 20 machine hours in order to still produce y_0 units of output. It can be shown that:[26]

$$MRTS \equiv -(\Delta K/\Delta L)y_0 = MP_L/MP_K > 0 \qquad (26.10)$$

where $(\Delta K/\Delta L)y_0$ is the slope of the particular isoquant y_0, MP_K is the marginal product of capital, and MP_L the marginal product of labor.

Equation 26.10 is valuable because it links the isoquant's slope to the marginal products of labor and capital. Remember, because of the law of diminishing marginal returns, the MP_L declines as L rises, and the MP_K declines as K rises—all else equal. Heuristically, at point C capital is plentiful (so the MP_K is small) and labor is scarce (so the MP_L is large). But, in Equation 26.10, a large number divided by a small one is still a large number, so the MRTS is large (and the isoquant is steep). Similar but opposite considerations apply to point A, where, this time, labor is the abundant factor, and the slope of the isoquant is relatively flat.

Isocost Lines. As the name perhaps already suggests (*iso* = "equal" and *cost* = "cost") *isocost* means "equal cost." In this context, an **isocost line** (or curve) describes the locus of equally costly capital and labor input combinations (or *bundles*, as they are sometimes called). Figure 26.9*a* depicts a family of isocost lines. For example, the line C_0 represents the locus of K, L bundles that cost the firm exactly C_0 dollars. Note that the isocost lines are parallel and that the firm's total costs decrease in the direction indicated by the arrow, implying $C_0 < C_1$. As we will now see, none of these facts should, however, be the cause for much surprise.

FIGURE 26.9   A Family of Isocost Lines
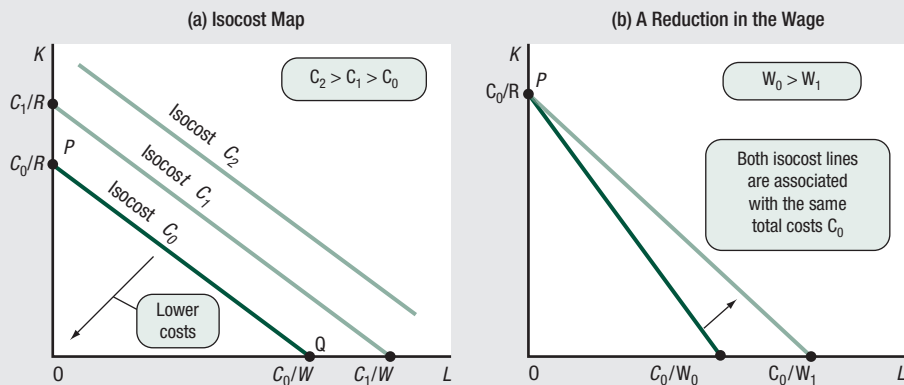
Panel (a) depicts a family of isocost lines. Each of the lines is parallel, and total costs decline in the direction indicated by the arrow. Panel (b) depicts the effects on a particular isocost line of a reduction in the wage rate. Notice that the effect is that the isocost line pivots outward around point P.

By definition, the firm's total costs are $C = wL + RK$. If we treat C as given and solve for K, the result is the general equation for an *isocost line*,

$$K = (C/R) - (W/R) \cdot L \qquad (26.11)$$

This equation determines the number of machine hours the firm can afford to hire given its total costs are C and given it hires L labor hours. The reader may recognize it as the equation of a straight line that has a slope of $-(W/R)$, a vertical intercept of (C/R), and a horizontal intercept of (C/W). The fact that all of the isocost lines have the same slope, $-(W/R)$, explains why the isocost map depicted in Figure 26.9b consists of a series of parallel straight lines.

The vertical intercept, C/R, equals the maximum amount of capital the firm can afford to hire given the rental price, R, and its willingness to spend $\$C$. Similarly, the horizontal intercept, C/W, equals the maximum amount of labor the firm can hire given W and C. Clearly, the maximum amount of labor or capital the firm can afford increases with the amount it is willing to spend, which explains why $C_1/W > C_0/W$ and $C_1/R > C_0/R$.

Finally, Figure 26.10b shows that a *ceteris paribus* reduction in the wage, from W_0 to W_1, causes the isocost line C_0 to pivot outward around point P, thus becoming shallower in absolute value. The vertical intercept, C_0/R, equals the maximum amount of capital the firm can hire given C_0 and R. Consequently, point P does not budge—and the isocost line pivots—because C_0/R does not depend on W. However, given C_0, the firm can afford to hire more labor as the wage declines,

$C_0/W_1 > C_0/W_0$, which explains the outward shift.

The Cost-Minimizing Input Bundle. We now have all of the ingredients required to determine the cost-minimizing technique of producing any target level of output, y. The key to our success is to remember, as shown in Figure 26.9a, that the firm's total costs are lowest when it selects input bundles that lie as close as possible to the origin.

Figure 26.10 explains the basic principles. To begin with, let us suppose that the manager's goal is to minimize the costs of producing the particular target level of output y_0. As shown, the isoquant labeled y_0 circumscribes the locus of input combinations,

FIGURE 26.10   Cost Minimization

K,L, that he can call on to meet his production target. Nevertheless, these combinations differ in their costs, and the manager's goal is to determine the particular one that minimizes them.

Suppose that the manager selects the input bundle labeled I, which lies on the isocost line C_{-1}. The firm's total costs C_{-1} are relatively low because this isocost line is close to the origin. The snag, however, is that if the manager chooses this input bundle, then he will fail to meet his production target: point I does not lie on the desired isoquant y_0. Indeed, it is clear that all of the input combinations that lie along the isocost line C_{-1} must then be excluded on these grounds. Now suppose that he picks point F. This option is feasible because it does at least lie on the requisite isoquant y_0. Nevertheless, it is not optimal, in the sense of minimizing the firm's costs. To see why, notice point Q is also feasible and it is less costly: $C_0 < C_1$. In fact it is relatively easy to see that point Q—with the associated input levels K_0^* and L_0^*—represents the optimal cost-minimizing choice. Any lower-cost input combination (such as I) is infeasible, and any other feasible combination (such as point F) is more costly.

The cost-minimizing input bundle, Q, is located at the point of tangency between the isoquant, y_0, and the isocost line, C_0, so both curves touch and share the same slope. In fact, after running through essentially the same steps, the cost-minimizing input choices associated with any given target output level y necessarily occurs at the point of tangency between the relevant isoquant and an isocost line.

For example, the least-cost method of producing y_1 is located at point R in the figure, and the least-cost method of producing y_2 at point S. Holding fixed the input prices W and R, the **output-expansion path**, XX, traces out the locus of tangency points (and hence the cost-minimizing input bundles) associated with
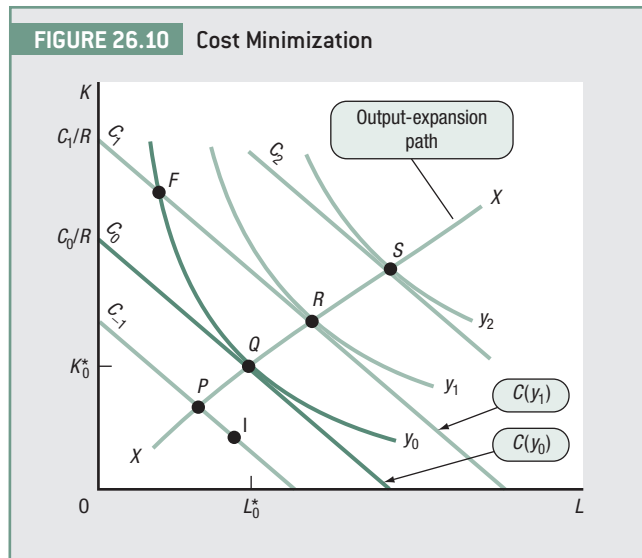
each level of output y. Under the plausible assumption that capital and labor are **normal inputs**, which means the firm demands more of them following a *ceteris paribus* increase in its output, the slope of the output-expansion path, XX, is necessarily *positive*.

The slope of a given isoquant is $-MP_L/MP_K$ (Equation 26.10), and the slope of an isocost line is $-(W/R)$ (Equation 26.11). Along the output-expansion path the two slopes are equal. Canceling the minus signs, this implies $W/R = MP_L/MP_K$, which can be rearranged as:

$$MP_L/W = MP_K/R \qquad (26.12)$$

But this is the equal-bang-for-the-buck condition we derived when we studied the input approach—see Equation 26.3. In other words, this condition is really telling us that the profit-maximizing firm optimally chooses its input levels in the most cost-effective manner. Worked Problem 26.5 shows how to find the firm's cost-minimizing input bundle and characterize its properties.

The Cost Function. At a given point of tangency between an isoquant and an isocost line, the two curves not only share the same slope but they also touch. For example, point Q lies on the isocost line C_0. Hence the minimum cost of producing y_0 units of output, denoted by $C(y_0)$, is $C(y_0) = C_0$. Likewise, the lowest-cost

---

Worked Problem 26.5

Determining the Cost-minimizing Input Bundle

Problem. A particular firm possesses the production function $y = \sqrt{K} \cdot \sqrt{L}$.

(a) What is the least cost method for it to produce 100 units of output if $W = \$64$ and $R = \$4$?

(b) What happens if the rental price of capital increases to $R = \$16$?

Hint. It can be shown, $MP_K = 0.5 \sqrt{L}/\sqrt{K}$ and $MP_L = 0.5 \sqrt{K}/\sqrt{L}$.

Solution. The cost-minimizing input bundle is located at the point of tangency between the particular isoquant $y = 100$ and the lowest possible isocost line. At this point, both lines share the same slope, so $W/R = MRTS$. Using Equation 26.10 and the hint, it then follows $MRTS = MP_L/MP_K = K/L = W/R$. Solving for K yields $K = (W/R) \cdot L$. Using this

result, $y = 100$, and the production function implies $100 = \sqrt{(W/R)} \cdot L$. Rearranging and solving for L gives the optimal (*) solution: $L^* = 100 \cdot \sqrt{(R/W)}$. Finally, using this result and $K = (W/R) \cdot L$ yields $K^* = 100 \cdot \sqrt{(W/R)}$.

(a) By using the above solutions for the cost-minimizing inputs, it follows that if $W = \$64$ and $R = \$4$, then $L_0^* = 100 \cdot \sqrt{(1/16)} = 25$ and $K_0^* = 100 \cdot \sqrt{16} = 400$. The firm's total costs are, $C_0 = 400(4) + 25 \cdot (64) = \$3200$.

(b) If the price of capital increases to $R = \$16$, then $L_1^* = 100\sqrt{16/64} = 50$ and $K_1^* = 100\sqrt{4} = 200$. By comparing the findings in (a) and (b), notice that the firm responds to the increase in R by substituting toward labor and away from the more expensive capital. Its total costs also increase: $C_1 = \$6400$.

method of producing y_1 is located at point R in the figure, and it is associated with a total cost of $C(y_1) = C_1$.

Together, these considerations allow us to derive the firm's **cost function**, denoted $C(y)$, which determines the minimum total cost of producing any given output level y. Some of the major properties of the cost function, $C(y)$, are obvious enough straight off the bat. In particular, $C(0) = 0$ because a nonexistent phantom firm, which produces nothing at all, obviously incurs no costs. Moreover, the firm's total costs, $C(y)$, increase with y because the firm must pay for the extra inputs that are required to produce the greater level of output.

The reader may recall that **marginal cost,** denoted $MC(y)$, equals the increase in the firm's total costs following a unit increase in its output. More specifically, let Δy denote a small change in output and let ΔC denote the resulting change in the firm's total costs, then marginal cost is defined by $MC(y) = \Delta C/\Delta y$.

FIGURE 26.11   The Optimal Output Level

Figure 26.11 depicts the hypothetical marginal cost schedule, $MC(y)_0$, that is pertinent to, say, Betsy's pizza business. (For the moment, disregard the other lines and curves in the figure.) As shown at point A, if she produces $y = 200$ pizzas, then $MC(y)_0 = \$10$ per pizza. This says that if she makes another pizza (for a grand total of 201) then her costs would increase by $10. Notice that her firm's marginal costs are not only positive, but they are also assumed to increase with the level of output y. The fact that they are positive tells us that her firm's total production costs increase with y. The positive slope of the $MC(y)_0$ curve tells us more: it is increasingly difficult and costly for her firm to produce greater and greater quantities of output.

It is readily seen why this assumption is plausible in the case of her pizza parlor and by extension why it is plausible for many other businesses as well. Imagine that Betsy's managerial talents are well suited to running a small restaurant that produces 200 pizzas a day but not to running a larger one that produces 400 a day. At point A, she is in her managerial comfort zone and her marginal costs are only $MC(200)_0 = \$10$ per pizza. Matters are, however, quite different at point B because her lack of experience potentially results in her making many poor decisions (and here one has to think only of a daily assortment of either burnt or undercooked pizzas). The resulting waste explains the much higher marginal cost of $MC(400)_0 = \$15$ per pizza and, in turn, the higher marginal cost accounts for the positive slope of the $MC(y)_0$ schedule.

Finally, Figure 26.11 depicts the effects of a reduction in either the wage, W, or the rental price of capital, R, on her firm's marginal costs, which causes the marginal-cost curve to shift rightward from MC_0 to MC_1. This indicates that a reduction in one or more input prices translates into lower total and marginal costs.
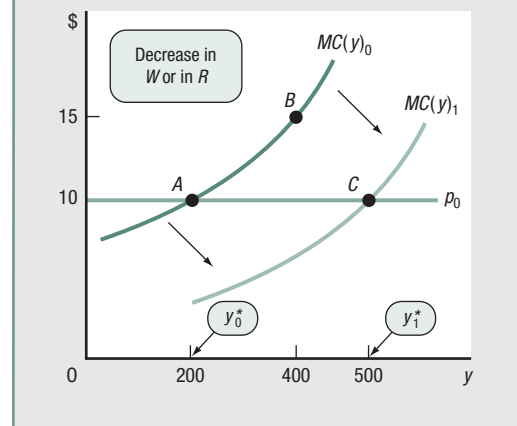
Step 2: Profit Maximization

As shown on page 31, by using the cost function $C(y)$, the firm's profits can be written $\Pi(y) = py - C(y)$. It follows that the manager's problem boils down just to discovering the output level, denoted y^*, that maximizes them. This endeavor is straightforward if he uses the one-step-at-a-time principle discussed in Chapter 3. Major Result 26.5 summarizes the key condition that governs the firm's optimal behavior, and the explanation and interpretation follow.

MAJOR RESULT 26.5

The Output Approach

The optimal profit-maximizing level of output, y^*, is governed by:

$$p = MC\,(y^*) \tag{26.13}$$

By definition marginal revenue, MR, equals the increase in the firm's revenues following the production and sale of an additional unit of output. Under competitive conditions, $MR = p$ because the firm is a price taker and consequently each additional sale increases its revenues by precisely $\$p$. Viewed in this light, Equation 26.13 simply says that the optimal level of output, y^*, is governed by the perhaps already familiar condition that marginal revenue must equal marginal cost.

To better understand this result, suppose Betsy's marginal costs are given by the curve $MC(y)_0$ in Figure 26.11 and that she is currently producing, say, 400 pizzas a day. Another quick glance at the figure shows $\$10 = p_0 < MC\,(400)_0 = \$15$. However, under these circumstances, if she takes a single step and reduces her output by one pizza, then she can increase her profits by $\$(15 - 5) > 0$. This implies producing 400 pizzas cannot be optimal because the maximum cannot be improved on. In other words, the condition $p - MC(y) \neq 0$ is akin to a green light that tells her to keep adjusting her output level because opportunities still remain for her to increase the profits. In contrast, the condition $p - MC(y) = 0$ is a red light that tells her to stop adjusting her output level because there is no further chance for improvement.

In Figure 26.11, the firm's initial optimal choice is located at the intersection point, A, which implies its profit-maximizing output level is $y_0^* = 200$ pizzas. The figure also depicts the effects of a decrease in the wage (or the rental price of capital), which shifts the firm's marginal-cost curve rightward to MC_1. The new intersection point is located at C, indicating that it is now optimal for it to produce $y_1^* = 500$ pizzas. Hence a reduction in one or more input price leads to an unambiguous increase in the firm's profit-maximizing level of output.

The Long-Run Demand for Labor

With a commendable display of persistence, let us now examine how the two steps described above can be used to shed light on the nature of the firm's long-run

demand for labor. To this end, Figure 26.12 shows how a firm optimally responds to a reduction in the wage, from W_0 to W_1.

Given the initial wage, W_0, the firm's profit-maximizing level of output is y_0^*. The initial cost-minimizing input bundle is located at point E: the point of tangency between the isoquant y_0^* and the isocost line C_0. Hence, given W_0, the firm's optimal behavior is described by the following facts: it produces y_0^* units of output, it demands L_0^* labor hours, and it demands K_0^* capital hours.

The reduction in the wage from W_0 to W_1 lowers the firm's marginal costs, which encourages it to expand its profit-maximizing output level from y_0^* to y_1^* (see Figure 26.11 and the associated discussion). Moreover, it

FIGURE 26.12   The Effects of a Reduction in the Wage

causes each of the isocost lines to become shallower in absolute value, which can be seen by comparing the slopes of the two isocost lines. The result is that the optimal cost-minimizing input bundle shifts from point E to point F—the point of tangency between the new target isoquant y_1^* and the shallower isocost line C_1. Hence the firm responds to the wage cut by demanding more labor $(L_1^* > L_0^*)$ and more capital $(K_1^* > K_0^*)$. This finding indicates that the firm's long-run labor-demand curve is negatively sloped. It also suggests that labor and capital are gross complements because a reduction in the wage increases the demand for labor and capital.

Yet, before we all reach for our celebratory party hats and kazoos, we had better check that these findings are robust, in the sense they extend beyond the particular firm depicted in Figure 26.12. On this score, point G sends an apparently ominous warning sign. Think of the isoquant $y_1^{*\prime}$ as belonging to a *different* firm that shares the same initial isoquant $y_0^{*\prime}$. Notice that this isoquant is perfectly respectable (in the sense that it satisfies all of the properties laid out on page 32) but that the point of tangency, G, is located on the isocost line C_1 below K_0^*. Hence, as far as this firm is concerned, capital and labor are gross substitutes: a reduction in the wage increases its demand for labor but lowers its demand for capital. (Remember: Complements go together and substitutes go in opposite directions.)

The Scale and Substitution Effects

To claim that a reduction in the wage might induce a firm to hire more capital, less capital, or, for that matter, even the same amount of capital would appear to claim little at all. Fortunately, however, this is not the case because the underlying reasons for the theoretical ambiguity are of fundamental
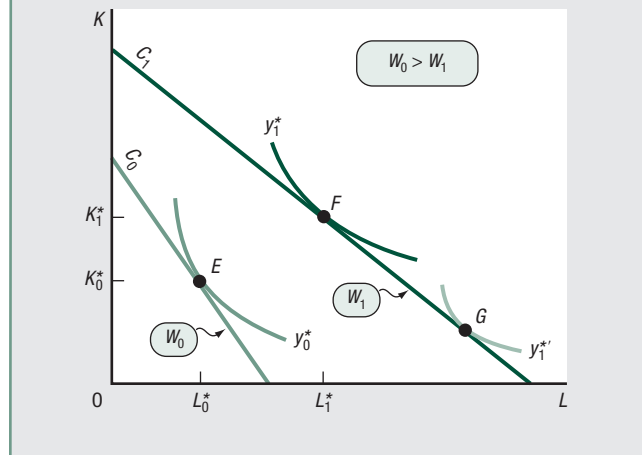
> **Tip!**
> *Don't forget that following the reduction in the wage, the optimal level of output increases. Hence the isocost line both pivots and shifts outward.*

importance, and economic theory allows us to say something sensible about them. More specifically, the reduction in the wage unleashes the following two major economic forces:

- **The Scale Effect.** Holding constant the wage at its initial value, W_0, the expansion in output from y_0^* to y_1^* encourages the firm to hire more labor and more capital (at least under the plausible assumption that these two inputs are normal).

- **The Substitution Effect.** Holding constant the level of output at its new value, y_1^*, the reduction in the relative price of labor (and associated increase in the relative price of capital) encourages the firm to switch to the less costly alternative and so hire more labor and less capital.

Major Result 26.6 summarizes the main implications of these findings for the firm's optimal long-run behavior.


MAJOR RESULT 26.6

The Firm's Long-Run Behavior

Following a reduction in the wage:

(a) There is an unambiguous increase in the long-run demand for labor because the scale and substitution effects work in harmony.

(b) If the scale effect dominates the substitution effect, then the firm's demand for capital also increases, which implies capital and labor are gross complements.

(c) If the substitution effect dominates the scale effect, then the firm's demand for capital decreases, which implies labor and capital are gross substitutes.

The Scale and Substitution Effects: A Graphical Approach. Figure 26.13 reproduces the main facts presented in Figure 26.12 so as to provide a graphical depiction of the scale and substitution effects. The figure conveys the following message: in the journey that takes the firm from its point of embarkation, E, to its final destination, F, the figure shows it stopping off at a point of interest Q. The trick is that we have arranged its travel plans so that the journey from $E \rightarrow Q$ captures the scale effect, and the continuation $Q \rightarrow F$ the substitution effect.

Bearing this travel itinerary in mind, suppose that the wage remains unchanged at W_0 but that the firm expands its output level from y_0^* to its new optimal level y_1^*. Given the initial wage W_0 and the price of capital R, the least-cost method of producing y_1^* is located at the point Q in the figure: the point of tangency between the new destination isoquant and the isocost line SS, which is parallel to the original C_0. Notice that point Q necessarily lies on the positively sloped output-expansion path XX because the expansion path traces out all such tangency points. It follows that, along this leg of the journey, there is an unambiguous increase in the firm's demand for capital and labor: $K_{sc}^* > K_0^*$ and $L_{sc}^* > L_0^*$, where sc stands for scale effect.

FIGURE 26.13   The Scale and Substitution Effects
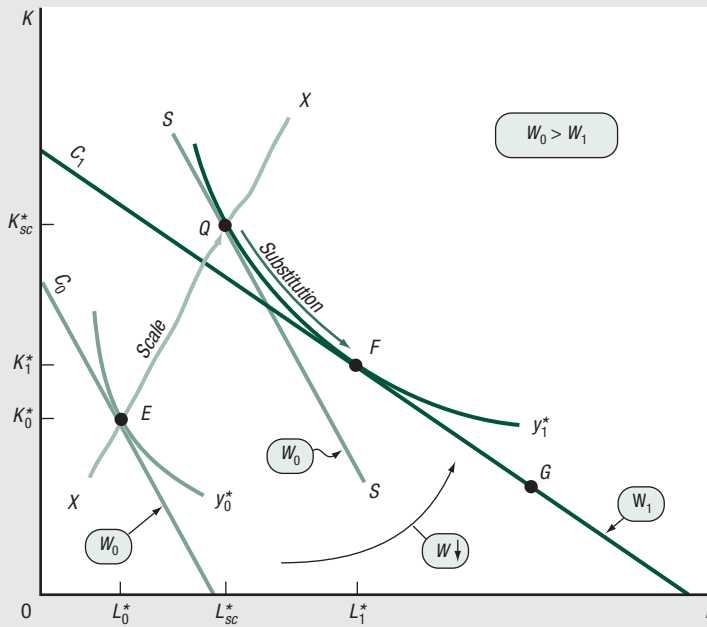
The figure captures **scale** and **substitution** effects resulting from the reduction in the wage from W_0 to W_1. The firm initially produces y_0^* units of output, and its optimal input bundle is located at the tangency point E. After the decrease in the wage, its optimal output level increases to y_1^*, and its cost-minimizing input bundle shifts to F.

The scale effect (sc) is captured by the movement from E to Q. It isolates the impact of the increase in output on the firm's input demands—at the initial input prices. Notice the new cost-minimizing bundle lies on the positively sloped output-expansion path XX, so the scale effect leads to an unambiguous increase in the demand for labor and capital.

The substitution effect is captured by the movement from point Q to F. It isolates the impact of the relative change in input prices on the firm's behavior. Because isoquants are negatively sloped, the substitution effect leads to an increase in the firm's demand for labor and reduction in its demand for capital.

For the second stage of the journey, imagine that the firm begins at point Q and that the wage is gradually lowered from W_0 to W_1—while all the time its output is held constant at y_1^*. This change causes the isocost line to swivel around the isoquant y_1^* from SS to C_1, and the cost-minimizing input bundle to shift from its temporary resting stop Q to its final destination F.

Clearly there is no scale effect: the level of output is the same before and after the change. Instead, the movement from point Q to F captures the substitution effect and thus isolates the firm's response to the relative increase in the price of capital (or, equivalently, the relative reduction in the price of labor). The fact that the firm's production isoquants are negatively sloped implies that the tangency point, F, must lie to the left of point Q. Consequently, the substitution effect results in an unambiguous increase in the firm's demand for labor and decrease in its demand for capital: $K_{sc}^* > K_1^*$ and $L_1^* > L_{sc}^*$.[27]

Together, the scale and substitution effects explain why a decrease in the wage unambiguously increases the firm's demand for labor but has a theoretically ambiguous influence on its demand for capital. Thus, as far as labor is concerned, the two effects work in harmony: the journey from E to Q (scale) and the continuation from Q to F (substitution) both raise the firm's demand for labor. In contrast, the scale and substitution effects have a conflicting effect on the firm's demand for

capital: the scale effect, $E \rightarrow Q$, encourages it to hire more capital, but the substitution effect, $Q \rightarrow F$, encourages it to hire less. The outcome of this battle determines whether it ultimately hires more capital (point F) or less (point G), and thus whether capital and labor are gross complements or gross substitutes.

## The Magnitude of the Substitution Effect

The magnitude of the substitution effect plays a central role in shaping a firm's optimal long-run behavior. It clearly influences how much the firm's demand for labor responds to a given change in the wage because the response itself equals the scale effect plus the substitution effect. Moreover, it is instrumental in determining whether capital and labor are gross substitutes or complements in production.

As we shall now see, the size of the substitution effect is related to certain fundamental characteristics of the firm's production technology. Figure 26.14 is used to explain the basic principles: it depicts three very different isoquant maps that correspond to three very different technologies.

Each of the three panels of Figure 26.14 depict the substitution effect resulting from a reduction in the wage. Below, we will discuss each case in turn.

- **The Standard Case (Regular Isoquants).** Figure 26.14*a* requires little comment. It is basically a stripped down version of Figure 26.13, in which an assortment of details have been suppressed. The isoquants have a regular shape and the substitution effect, induced by the reduction in the wage, is captured by the shift in the lowest-cost input bundle from point F to its final destination Q.
- **Perfect Substitutes.** The isoquant map shown in Figure 26.14*b* is noteworthy because it consists of a series of negatively sloped straight lines. One conse-

---

FIGURE 26.14 The Size of the Substitution Effect Depends on the Curvature of the Firm's Isoquants

(a) Regular Isoquants (b) Perfect Substitutes (c) Perfect Complements
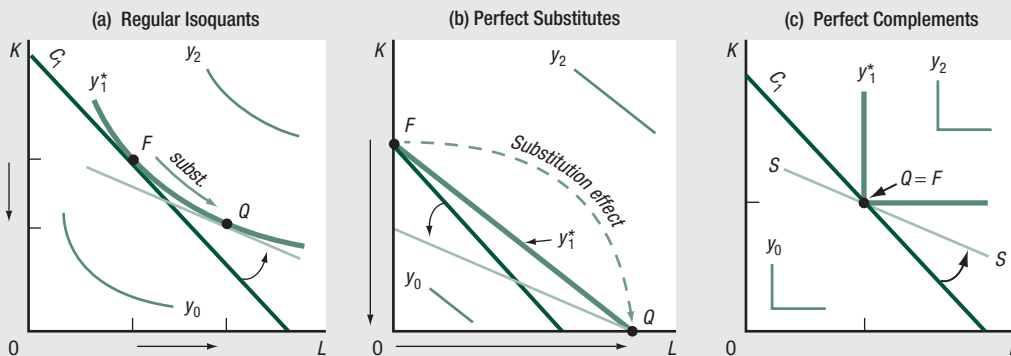
quence of this is that the firm can produce y_1^* using exclusively capital (point F), labor (point Q), or any mixture of the two that lies along the isoquant FQ.

Prior to the reduction in the wage, the least-cost method of producing y_1^* is located at point F in the figure (why?). The reduction in the wage causes the isocost lines to become shallower in absolute value. The cost-minimizing input bundle jumps from point F to point Q which, in turn, unleashes a massive substitution effect that is captured by the dashed arrow. In deference to the huge size of the effect—it could not be any larger—labor and capital are called perfect substitutes.

Generically, two inputs are perfect substitutes only if they are essentially identical. (For this reason it is unlikely that capital and labor would ever be perfect substitutes because they are so fundamentally different in kind.) For example, suppose that a firm uses steel rolls in its manufacturing process, and it purchases K square meters of steel from one supplier and L square yards of (identical) steel from another. It is easy to see that, in this case, K and L would in fact be perfect substitutes because all the firm really cares about is the total quantity of steel measured in a common unit.

- **Perfect Complements.** At first glance, Figure 26.14c is quite startling because the isoquants are L-shaped.[28] It follows from their shape that an increase in capital or labor alone has no effect whatsoever on the firm's output. Instead, the firm can raise its output level only if it simultaneously employs more of both inputs.

Once again, prior to the reduction in the wage, the least-cost method of producing y_1^* is located at point Q in the figure—remember the firm is trying to get the isocost line SS as close to the origin as possible while maintaining contact with the isoquant. After the reduction in the wage, the isocost lines become shallower in absolute value (which can be seen by comparing the slopes of the isocost lines SS and C_1). Nevertheless, it is easy to see that the least-cost method of producing y_1^* does not budge at all following the change. This implies that $Q = F$ and that the substitution effect is zero. In a pleasantly suggestive way, capital and labor are termed perfect complements because of the complete absence of the substitution effect.

Just as was the case of perfect substitutability, the concept of perfect complementarity is an idealized limiting case. Nevertheless, the following simple illustration helps to clarify the circumstances under which it may seem reasonable.

In Figure 26.14c let K denote buses, L bus drivers, and y the number of passengers carried per day. Buses do not drive themselves, so each of them needs a driver. It follows that the bus company's isoquants are L-shaped, because the number of passengers carried per day remains unchanged if it increases the number of buses (holding the number of drivers constant) or it increases the number of drivers (holding the number of buses constant). In fact, in order to carry more passengers, the company clearly must increase both the number of buses *and* the number of drivers.[29]

Technical Appendix 26.B The Mathematical Approach

In this appendix, several of the results presented mathematically in the chapter are derived. Throughout, familiarity with the notation developed in Appendix 3.A is assumed.

In the long run the firm can adjust its employment level and its capital stock. Its technology is given by $y = F(K, L)$, which is assumed to be increasing in K and L and to be strictly concave. A basic result in mathematics is that the concavity of $F(K, L)$ implies that:

$$F_{LL} < 0, \quad F_{KK} < 0 \qquad (26.14)$$

$$F_{LL} F_{KK} - (F_{KL})^2 > 0 \qquad (26.15)$$

where (using our standard convention) $F_{LL} \equiv \partial^2 y / \partial L^2$ and likewise for the other second-order partial derivatives. We will use this result when examining the properties of the firm's long-run demand for labor.

As shown on page 44, by using the expression $y = F(K, L)$ to substitute out y, the firm's profits can be written:

$$\Pi(K, L) = pF(KL) - (WL + RK) \qquad (26.16)$$

The manager's goal is to discover the optimal input levels—denoted K^* and L^*—that maximize the firm's profits. The first-order conditions required for a maximum are:

$$pF_L - W \equiv 0 \quad \text{and} \quad pF_K - R \equiv 0 \qquad (26.17)$$

where $F_L \equiv MP_L = \partial y / \partial L$ is the marginal product of labor, and analogously F_K is the marginal product of capital. It is understood that Equations 26.17 are evaluated at the optimal values of K and L.

As any mathematician will happily testify, it is one thing to have a pair of equations, like those in 26.17, and quite another to have the faintest clue about what to do with them. One approach is to bring more information to bear on the problem (by, for example, specifying a particular formula for $y = F(K, L)$), and another is to mathematically analyze the equations in the hope of establishing some general results. We begin by pursuing the first approach because it is somewhat simpler.

The Cobb-Douglas Production Technology. As noted in Chapter 3, because of its simplicity and tractability, the Cobb-Douglas production function is often deemed to be the gold standard of production functions. Suppose then, that the firm's technology can be represented by the following Cobb-Douglas form:

$$y = K^\alpha L^\beta, \quad \text{where} \quad \alpha, \beta > 0 \quad \text{and} \quad \alpha + \beta < 1 \qquad (26.18)$$

The plan of attack is that we are going to apply the general result given by Equations 26.17 to the particular case of the Cobb-Douglas production technology.

To do so, it is necessary to evaluate the partial derivatives F_L and F_K. It is readily checked that,

$$F_K = \alpha y / K \quad \text{and} \quad F_L = \beta y / L \tag{26.19}$$

Using these terms in Equation 26.17 yields, $\alpha p y / K = R$ and $\beta p y / L = W$. In order to make further progress, it is convenient to set $\hat{w} \equiv W(\beta \rho)$ and $\hat{r} = (R / \alpha \rho)$. After doing so, we have $y / K = \hat{r}$ and $y / L = \hat{w}$. Simple rearrangement gives,

$$(\hat{w}/\hat{r}) = K/L \quad \text{and} \quad (\hat{r}/\hat{w}) = L/K \tag{26.20}$$

The first-order conditions can be written as,

$$\hat{w} = (K/L)^\alpha L^{(\alpha + \beta) - 1} \tag{26.21}$$

$$\hat{r} = K^{(\alpha + \beta) - 1} (L/K)^\beta \tag{26.22}$$

Substituting for (K/L) in the first and (L/K) in the second gives,

$$\hat{w} = (\hat{w}/\hat{r})^\alpha L^{(\alpha + \beta) - 1} \tag{26.23}$$

$$\hat{r} = (\hat{r}/\hat{w})^\beta K^{(\alpha + \beta) - 1} \tag{26.24}$$

Solving these equations for K and L yields the explicit long-run demand functions for capital and labor,

$$L = \hat{w}^{-\frac{1 - \beta}{1 - (\alpha + \beta)}} \cdot \hat{r}^{-\frac{\alpha}{1 - (\alpha + \beta)}} \tag{26.25}$$

$$K = \hat{w}^{-\frac{\beta}{1 - (\alpha + \beta)}} \cdot \hat{r}^{-\frac{1 - \alpha}{1 - (\alpha + \beta)}} \tag{26.26}$$

Notice the pleasing symmetry that exists between them.

Furthermore, because $1 - (\alpha + \beta) > 0$ it is readily seen that an increase in W lowers the demand for labor and that an increase in R lowers the demand for capital—so the respective demand curves are downward sloping. Furthermore, capital and labor are gross complements because an increase in W reduces K, and an increase in R reduces L.

The General Case. If we let $w = W/p$ and $r = R/p$ denote the *real wage* and the *real rental price* respectively, then the first-order conditions take the simple form,

$$F_L - w \equiv 0 \quad \text{and} \quad F_K - r \equiv 0 \tag{26.27}$$

Can we use these equations to say anything sensible about the properties of the long-run demand for labor? Remarkably, the answer is yes because of the following five facts: $F_L > 0$, $F_K > 0$, $F_{KK} < 0$, $F_{LL} < 0$, and $\Delta \equiv F_{LL} F_{KK} - (F_{KL})^2 > 0$. The first two simply assert that capital and labor are productive inputs (so the more of them the merrier); the second two, that the technology exhibits diminishing marginal returns to capital and labor; and the last one, that "average" input bundles are better than extreme ones. (In fact, the last three facts follow from the concavity of $F(K, L)$.)

If the reader ever happens to be confronted with analyzing this sort of problem, the method of attack is always basically the same. First, totally differentiate the system. Second, write out the total derivatives. Finally, solve for them (often using Cramer's rule). An example illustrates the method. Suppose that we want to determine how the firm's long-run demand for labor and capital are affected by an increase in the (real) wage w. To investigate this problem, it is necessary that we evaluate the total derivatives dL/dw and dK/dw. With this goal in mind, totally differentiate the two first-order conditions:

$$F_{LL}dL + F_{KL}dK - dw = 0 \qquad (26.28)$$

$$F_{KL}dL + F_{KK}dK - dr = 0 \qquad (26.29)$$

where the fact that $F_{KL} = F_{LK}$ is used, which follows from Young's theorem. Our goal is to evaluate the effect of a change in w, holding r constant. Therefore, set $dr = 0$ and divide by $dw \neq 0$. The result is,

$$F_{LL}(dL/dw) + F_{KL}(dK/dw) = 1 \qquad (26.30)$$

$$F_{KL}(dL/dw) + F_{KK}(dK/dw) = 0 \qquad (26.31)$$

In many respects this is a very simple equation system with two unknowns: dL/dw and dK/dw. The solution is,

$$dL/dw = F_{KK}/\Delta < 0 \qquad (26.32)$$

$$dK/dw = -F_{KL}/\Delta \qquad (26.33)$$

where $\Delta \equiv F_{LL}F_{KK} - (F_{KL})^2 > 0$. It follows from the fact that $F_{KK} < 0$—that is, there are diminishing returns to capital—the long-run labor-demand curve is necessarily downward sloping: $dL/dw < 0$.

Because $\Delta > 0$, the sign of dK/dw hinges on (and is the same as) the sign of the cross derivative F_{KL}. It is useful to recall that $F_{KL} = \partial F_K/\partial L$, implying F_{KL} describes how the marginal product of capital, F_K, responds to a small increase in L. From this vantage point, interpreting the result is now quite straightforward. Following the increase in the wage, the firm reduces its employment level. If $F_{KL} > 0$, then the reduction in L also lowers the productivity of capital and, as a consequence, the firm's demand for it. This implies that labor and capital are gross complements because $dK/dw < 0$. Similar, but opposite, remarks apply if $F_{KL} < 0$, in which case capital and labor are gross substitutes because $dK/dw > 0$.