READING

14

Topics in Demand and Supply Analysis

by Richard V. Eastin, PhD, and Gary L. Arbogast, PhD, CFA

Richard V. Eastin, PhD, is at the University of Southern California (USA). Gary L. Arbogast, PhD, CFA (USA).

LEARNING OUTCOMES

Mastery	The candidate should be able to
☐	**a.** calculate and interpret price, income, and cross-price elasticities of demand and describe factors that affect each measure;
☐	**b.** compare substitution and income effects;
☐	**c.** distinguish between normal goods and inferior goods;
☐	**d.** describe the phenomenon of diminishing marginal returns;
☐	**e.** determine and interpret breakeven and shutdown points of production;
☐	**f.** describe how economies of scale and diseconomies of scale affect costs.

INTRODUCTION

1

In a general sense, *economics* is the study of production, distribution, and consumption and can be divided into two broad areas of study: macroeconomics and microeconomics. **Macroeconomics** deals with aggregate economic quantities, such as national output and national income, and is rooted in **microeconomics**, which deals with markets and decision making of individual economic units, including consumers and businesses. Microeconomics is a logical starting point for the study of economics.

Microeconomics classifies private economic units into two groups: consumers (or households) and firms. These two groups give rise, respectively, to the theory of the consumer and the theory of the firm as two branches of study. The *theory of the consumer* deals with consumption (the demand for goods and services) by utility-maximizing individuals (i.e., individuals who make decisions that maximize the satisfaction received from present and future consumption). The *theory of the firm* deals with the supply of goods and services by profit-maximizing firms.

It is expected that candidates will be familiar with the basic concepts of demand and supply. This material is covered in detail in the recommended prerequisite readings. In this reading, we will explore how buyers and sellers interact to determine transaction prices and quantities. The reading is organized as follows: Section 2 discusses the consumer or demand side of the market model, and Section 3 discusses the supply side of the consumer goods market, paying particular attention to the firm's costs. Section 4 provides a summary of key points in the reading.

2 DEMAND ANALYSIS: THE CONSUMER

The fundamental model of the private-enterprise economy is the demand and supply model of the market. In this section, we examine three important topics concerning the demand side of the model: (1) elasticities, (2) substitution and income effects, and (3) normal and inferior goods. The candidate is assumed to have a basic understanding of the demand and supply model and to understand how a market discovers the equilibrium price at which the quantity willingly demanded by consumers at that price is just equal to the quantity willingly supplied by firms. Here, we explore more deeply some of the concepts underlying the demand side of the model.

2.1 Demand Concepts

The quantity of a good that consumers are willing to buy depends on a number of different variables. Perhaps the most important of those variables is the item's own price. In general, economists believe that as the price of a good rises, buyers will choose to buy less of it, and as its price falls, they buy more. This opinion is so nearly universal that it has come to be called the **law of demand**.

Although a good's own price is important in determining consumers' willingness to purchase it, other variables also influence that decision. Consumers' incomes, their tastes and preferences, and the prices of other goods that serve as substitutes or complements are just a few of the other variables that influence consumers' demand for a product or service. Economists attempt to capture all these influences in a relationship called the **demand function**. (A function is a relationship that assigns a unique value to a dependent variable for any given set of values of a group of independent variables.)

Equation 1 is an example of a demand function. In Equation 1, we are saying, "The quantity demanded of good X depends on (is a function of) the price of good X, consumers' income, and the price of good Y":

$$Q_x^d = f\left(P_x, I, P_y\right) \quad (1)$$

where

Q_x^d = the quantity demanded of some good X (such as per household demand for gasoline in liters per month)

P_x = the price per unit of good X (such as € per liter)

I = consumers' income (as in €1,000s per household annually)

P_y = the price of another good, Y. (There can be many other goods, not just one, and they can be complements or substitutes.)

Often, economists use simple linear equations to approximate real-world demand and supply functions in relevant ranges. Equation 2 illustrates a hypothetical example of our function for gasoline demand:

$$Q_x^d = 84.5 - 6.39P_x + 0.25I - 2P_y \quad (2)$$

where the quantity of gasoline demanded $\left(Q_x^d\right)$ is a function of the price of a liter of gasoline (P_x), consumers' income in €1,000s (I), and the average price of an automobile in €1,000s (P_y).

The signs of the coefficients on gasoline price (negative) and consumers' income (positive) reflect the relationship between those variables and the quantity of gasoline consumed. The negative sign on average automobile price indicates that if automobiles go up in price, fewer will likely be purchased and driven; hence, less gasoline will be consumed. (As discussed later, such a relationship would indicate that gasoline and automobiles have a negative cross-price elasticity of demand and are thus complements.)

To continue our example, suppose that the price of gasoline (P_x) is €1.48 per liter, per household income (I) is €50,000, and the price of the average automobile (P_y) is €20,000. In this case, this function would predict that the per-household monthly demand for gasoline would be 47.54 liters, calculated as follows:

$$Q_x^d = 84.5 - 6.39(1.48) + 0.25(50) - 2(20) = 47.54$$

recalling that income and automobile prices are measured in thousands. Note that the sign on the "own-price" variable (P_x) is negative; thus, as the price of gasoline rises, per household consumption would decrease by 6.39 liters per month for every €1 increase in gas price. **Own price** is used by economists to underscore that the reference is to the price of a good itself and not the price of some other good.

In our example, there are three independent variables in the demand function and one dependent variable. If any one of the independent variables changes, so does the quantity demanded. It is often desirable to concentrate on the relationship between the dependent variable and just one of the independent variables at a time. To accomplish this goal, we can hold the other independent variables constant and rewrite the equation.

For example, to concentrate on the relationship between the quantity demanded of the good and its own price, P_x, we hold constant the values of income and the price of good Y. In our example, those values are 50 and 20, respectively. The equation would then be rewritten as

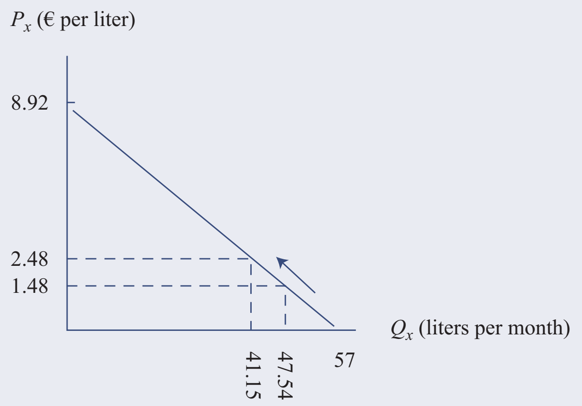
$$Q_x^d = 84.5 - 6.39P_x + 0.25(50) - 2(20) = 57 - 6.39P_x \quad (3)$$

The quantity of gasoline demanded is a function of the price of gasoline (6.39 per liter), per household income (€50,000), and the average price of an automobile (€20,000). Notice that income and the price of automobiles are not ignored; they are simply held constant, and they are "collected" in the new constant term, 57 [84.5 + (0.25)(50) – (2)(20)]. Notice also that we can solve for P_x in terms of Q_x^d by rearranging Equation 3, which gives us Equation 4:

$$P_x = 8.92 - 0.156Q_x^d \quad (4)$$

Equation 4 gives the price of gasoline as a function of the quantity of gasoline consumed per month and is referred to as the **inverse demand function**. Q_x in Equation 4 must be restricted to be less than or equal to 57 so that price is not negative. The graph of the inverse demand function is called the **demand curve** and is shown in Exhibit 1.[1]

Exhibit 1 Household Demand Curve for Gasoline



The demand curve represents the highest quantity willingly purchased at each price as well as the highest price willingly paid for each quantity. In this example, this household would be willing to purchase 47.54 liters of gasoline per month at a price of €1.48 per liter. If price were to rise to €2.48 per liter, the household would be willing to purchase only 41.15 liters per month.

This demand curve is drawn with price on the vertical axis and quantity on the horizontal axis. It can be correctly interpreted as specifying *either* the highest quantity a household would buy at a given price *or* the highest price it would be willing to pay for a given quantity. In our example, at a price of €1.48 per liter, households would each be willing to buy 47.54 liters per month. Alternatively, the highest price they would be willing to pay for 47.54 liters per month is €1.48 per liter. If the price were to rise by €1, households would reduce the quantity they each bought by 6.39 units, to 41.15 liters. The slope of the demand curve is measured as the change in price, P, divided by the change in quantity, Q ($\Delta P/\Delta Q$, where Δ stands for "the change in"). In this case, the slope of the demand curve is 1/–6.39, or –0.156.

The general model of demand and supply can be highly useful in understanding directional changes in prices and quantities that result from shifts in one curve or the other. Often, though, we need to measure how sensitive quantity demanded or supplied is to changes in the independent variables that affect them. This is the concept of **elasticity of demand** and **elasticity of supply**. Fundamentally, all elasticities are calculated in the same way: They are ratios of percentage changes. Let us begin with the sensitivity of quantity demanded to changes in the own price.

1 Following usual practice, we show linear demand curves intersecting the quantity axis at a price of zero. Real-world demand functions may be non-linear in some or all parts of their domain. Thus, linear demand functions in practical cases are approximations of the true demand function that are useful for a relevant range of values.

2.2 Own-Price Elasticity of Demand

In Equation 1, we expressed the quantity demanded of some good as a function of several variables, one of which was the price of the good itself (the good's "own-price").

In Equation 3, we introduced a hypothetical household demand function for gasoline, assuming that the household's income and the price of another good (automobiles) were held constant. That function was given by the simple linear expression $Q_x^d = 57 - 6.39P_x$. Using this expression, if we were asked how sensitive the quantity of gasoline demanded is to changes in price, we might say that whenever price changes by one unit, quantity changes by 6.39 units in the opposite direction; for example, if price were to rise by €1, quantity demanded would fall by 6.39 liters per month. The coefficient on the price variable (–6.39) could be the measure of sensitivity we are seeking.

There is a drawback associated with that measure, however. It is dependent on the units in which we measured Q and P. When we want to describe the sensitivity of demand, we need to recall the specific units in which Q and P were measured—liters per month and euros per liter—in our example. This relationship cannot readily be extrapolated to other units of measure—for example, gallons and dollars. Economists, therefore, prefer to use a gauge of sensitivity that does not depend on units of measure. That metric is called **elasticity**. Elasticity is a general measure of how sensitive one variable is to any other variable, and it is expressed as the ratio of percentage changes in each variable: $\%\Delta y/\%\Delta x$. In the case of **own-price elasticity of demand**, that measure is illustrated in Equation 5:

$$E_{p_x}^d = \frac{\%\Delta Q_x^d}{\%\Delta P_x} \tag{5}$$

This equation expresses the sensitivity of the quantity demanded to a change in price. $E_{p_x}^d$ is the good's own-price elasticity and is equal to the percentage change in quantity demanded divided by the percentage change in price. This measure is independent of the units in which quantity and price are measured. If quantity demanded falls by 8% when price rises by 10%, then the elasticity of demand is simply –0.8. It does not matter whether we are measuring quantity in gallons per week or liters per day, and it does not matter whether we measure price in dollars per gallon or euros per liter; 10% is 10%, and 8% is 8%. So the ratio of the first to the second is still –0.8.

We can expand Equation 5 algebraically by noting that the percentage change in any variable x is simply the change in x (Δx) divided by the level of x. So, we can rewrite Equation 5, using a few simple steps, as

$$E_{p_x}^d = \frac{\%\Delta Q_x^d}{\%\Delta P_x} = \frac{\dfrac{\Delta Q_x^d}{Q_x^d}}{\dfrac{\Delta P_x}{P_x}} = \left(\frac{\Delta Q_x^d}{\Delta P_x}\right)\left(\frac{P_x}{Q_x^d}\right) \tag{6}$$

To get a better idea of price elasticity, it might be helpful to illustrate using our hypothetical demand function: $Q_x^d = 57 - 6.39P_x$. When the relationship between two variables is linear, $\Delta Q_x^d / \Delta P_x$ is equal to the slope coefficient on P_x in the demand function. Thus, in our example, the elasticity of demand is –6.39 multiplied by the ratio of price to quantity. We need to choose a price at which to calculate the elasticity coefficient. Using our hypothetical original price of €1.48, we can find the quantity associated with that particular price by inserting 1.48 into the demand function as given in Equation 3:

$$Q = 57 - (6.39)(1.48) = 47.54$$

and we find that Q = 47.54 liters per month.

The result of our calculation is that at a price of 1.48, the elasticity of our market demand function is −6.39(1.48/47.54) = −0.2. How do we interpret that value? It means, simply, that when price equals 1.48, a 1% rise in price would result in a fall in quantity demanded of 0.2%.

In our example, when the price is €1.48 per liter, demand is not very sensitive to changes in price because a 1% rise in price would reduce quantity demanded by only 0.2%. In this case, we would say that demand is **inelastic**. To be precise, when the magnitude (ignoring algebraic sign) of the own-price elasticity coefficient has a value of less than one, demand is said to be inelastic. When that magnitude is greater than one, demand is said to be **elastic**. And when the elasticity coefficient is equal to negative one, demand is said to be **unit elastic**, or unitary elastic. Note that if the law of demand holds, own-price elasticity of demand will always be negative because a rise in price will be associated with a fall in quantity demanded, but it can be either elastic (very sensitive to a change in price) or inelastic (insensitive to a change in price). In our hypothetical example, suppose the price of gasoline was very high, say, €5 per liter. In this case, the elasticity coefficient would be −1.28:

$$Q = 57 - (6.39)(5) = 25.05$$

and

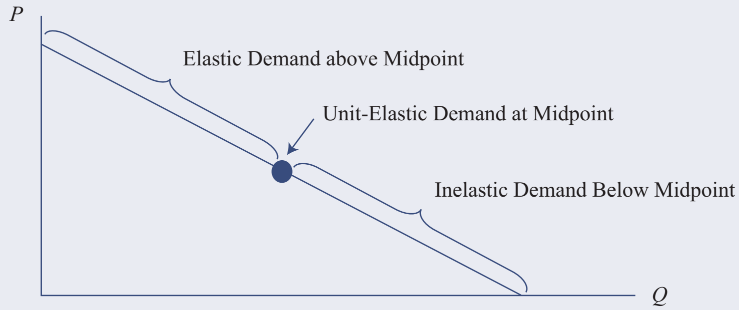
$$-6.39\ (5/25.05) = -1.28$$

Because the magnitude of the elasticity coefficient is greater than one, we know that demand is elastic at that price.[2] In other words, at lower prices (€1.48 per liter), a slight change in the price of gasoline does not have much effect on the quantity demanded, but when gasoline is expensive (€5 per liter), consumer demand for gas is highly affected by changes in price.

By examining Equation 6 more closely, we can see that for a linear demand curve the elasticity depends on where on the curve we calculate it. The first term, $\Delta Q/\Delta P$, which is the inverse of the slope of the demand curve, remains constant along the entire demand curve. But the second term, P/Q, changes depending on where we are on the demand curve. At very low prices, P/Q is very small, so demand is inelastic. But at very high prices, Q is low and P is high, so the ratio P/Q is very high and demand is elastic. Exhibit 2 illustrates a characteristic of all negatively sloped linear demand curves. Above the midpoint of the curve, demand is elastic; below the midpoint, demand is inelastic; and at the midpoint, demand is unit elastic.

2 If interested, evidence on price elasticities of demand for gasoline can be found in Molly Espey, "Explaining the Variation in Elasticity Estimates of Gasoline Demand in the United States: A Meta-analysis," *Energy Journal*, vol. 17, no. 3 (1996): 49–60. The robust estimates were about −0.26 for short-run elasticity—less than one year—and −0.58 for more than a year.

Exhibit 2 The Elasticity of a Linear Demand Curve

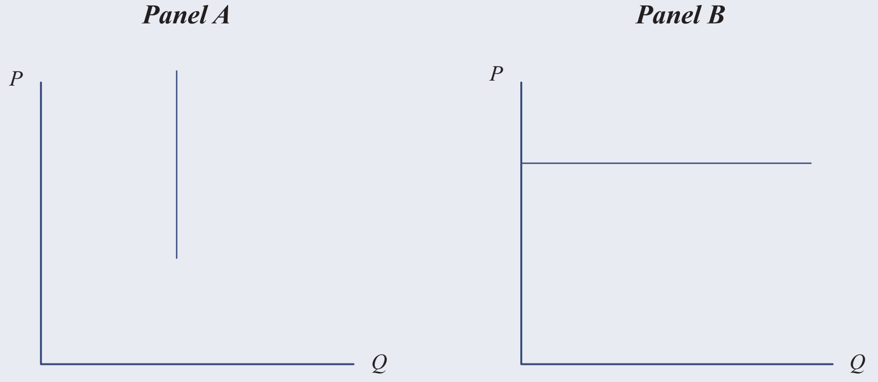


Note: For all negatively sloped, linear demand curves, elasticity varies depending on where it is calculated.

2.2.1 *Extremes of Price Elasticity*

There are two special cases in which linear demand curves have the same elasticity at all points: vertical demand curves and horizontal demand curves. Consider a vertical demand curve, as in Panel A of Exhibit 3, and a horizontal demand curve, as in Panel B. In the first case, the quantity demanded is the same, regardless of price. There is no demand curve that is perfectly vertical at all possible prices, but it is reasonable to assume that, over some range of prices, the same quantity would be purchased at a slightly higher price or a slightly lower price. Thus, in that price range, quantity demanded is not at all sensitive to price, and we would say that demand is **perfectly inelastic** in that range.

Exhibit 3 The Extremes of Price Elasticity



Note: A vertical demand has zero elasticity and is called perfectly inelastic.

Note: A horizontal demand has infinite elasticity and is called perfectly elastic.

In the second case, the demand curve is horizontal at some given price. It implies that even a minute price increase will reduce demand to zero, but at that given price, the consumer would buy some large, unknown amount. This situation is a reasonable description of the demand curve facing an individual seller in a perfectly competitive market, such as the wheat market. At the current market price of wheat, an individual farmer could sell all she has. If, however, she held out for a price above market price, it is reasonable to believe that she would not be able to sell any at all; other farmers' wheat is a perfect substitute for hers, so no one would be willing to buy any of hers at a higher price. In this case, we would say that the demand curve facing a seller under conditions of perfect competition is **perfectly elastic**.

2.2.2 *Predicting Demand Elasticity*

Own-price elasticity of demand is a measure of how sensitive the quantity demanded is to changes in the price of a good or service, but what characteristics of a good or its market might be informative in determining whether demand is highly elastic? Perhaps the most important characteristic is whether there are close substitutes for the good in question. If there are close substitutes for the good, then if its price rises even slightly, a consumer would tend to purchase much less of this good and switch to the less costly substitute. If there are no substitutes, however, then it is likely that the demand is much less elastic. Consider a consumer's demand for some broadly defined product, such as bread. There really are no close substitutes for the entire category of bread, which includes all types from French bread to pita bread to tortillas and so on. So, if the price of all bread were to rise, perhaps a consumer would purchase a little less of it each week, but probably not a significantly smaller amount. Now, consider that the consumer's demand is for a particular baker's specialty bread instead of the category "bread" as a whole. Surely, there are close substitutes for Baker Bob's Whole Wheat Bread with Sesame Seeds than for bread in general. We would expect, then, that the demand for Baker Bob's special loaf is much more elastic than for the entire category of bread.

In addition to the degree of substitutability, other characteristics tend to be generally predictive of a good's elasticity of demand. These include the portion of the typical budget that is spent on the good, the amount of time that is allowed to respond to the change in price, the extent to which the good is seen as necessary or optional, and so on. In general, if consumers tend to spend a very small portion of their budget on a good, their demand tends to be less elastic than if they spend a very large part of their income. Most people spend only a little on toothpaste each month, for example, so it really does not matter whether the price rises 10%. They would probably still buy about the same amount. If the price of housing were to rise significantly, however, most households would try to find a way to reduce the quantity they buy, at least in the long run.

This example leads to another characteristic regarding price elasticity. For most goods and services, the long-run demand is much more elastic than the short-run demand. For example, if the price of gasoline rises, we probably would not be able to respond quickly to reduce the quantity we consume. In the short run, we tend to be locked into modes of transportation, housing and employment location, and so on. With a longer adjustment period, however, we can adjust the quantity consumed in response to the change in price by adopting a new mode of transportation or reducing the distance of our commute. Hence, for most goods, long-run elasticity of demand is greater than short-run elasticity. Durable goods, however, tend to behave in the opposite way. If the price of washing machines were to fall, people might react quickly because they have an old machine that they know will need to be replaced fairly soon anyway. So when price falls, they might decide to go ahead and make a purchase. If the price of washing machines were to stay low forever, however, it is unlikely that a typical consumer would buy more machines over a lifetime.

Knowing whether the good or service is seen to be discretionary or non-discretionary helps to understand its sensitivity to a price change. Faced with the same percentage increase in prices, consumers are much more likely to give up their Friday night restaurant meal (discretionary) than they are to cut back significantly on staples in their pantry (non-discretionary). The more a good is seen as being necessary, the less elastic its demand is likely to be.

In summary, own-price elasticity of demand is likely to be greater (i.e., more sensitive) for items that have many close substitutes, occupy a large portion of the total budget, are seen to be optional instead of necessary, or have longer adjustment times. Obviously, not all these characteristics operate in the same direction for all goods, so elasticity is likely to be a complex result of these and other characteristics. In the end, the actual elasticity of demand for a particular good turns out to be an empirical fact that can be learned only from careful observation and, often, sophisticated statistical analysis.

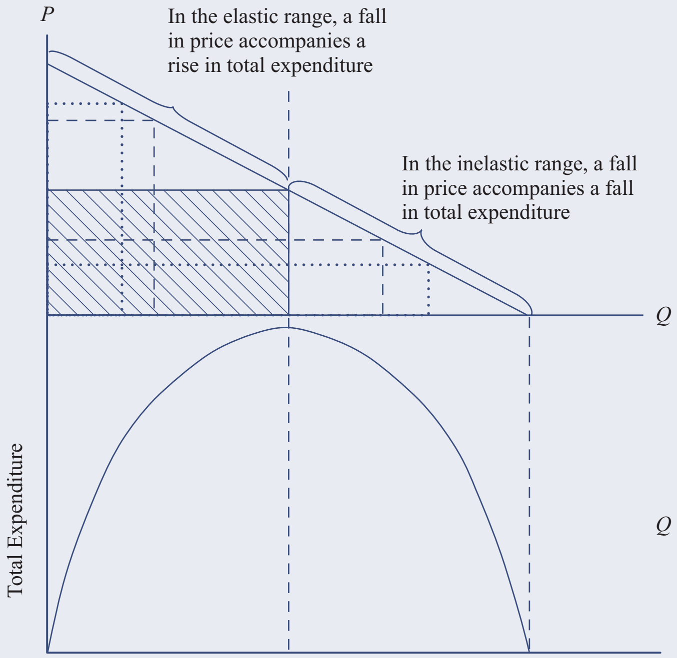
2.2.3 *Elasticity and Total Expenditure*

Because of the law of demand, an increase in price is associated with a decrease in the number of units demanded of some good or service. But what can we say about the total expenditure on that good? That is, what happens to price times quantity when price falls? Recall that elasticity is defined as the ratio of the percentage change in quantity demanded to the percentage change in price. So if demand is elastic, a decrease in price is associated with a larger percentage rise in quantity demanded. Although each unit of the good has a lower price, a sufficiently greater number of units are purchased so that total expenditure (price times quantity) would rise as price falls when demand is elastic.

If demand is inelastic, however, a given percentage decrease in price is associated with a smaller percentage rise in quantity demanded. Consequently, when demand is inelastic, a fall in price brings about a fall in total expenditure.

In summary, when demand is elastic, price and total expenditure move in *opposite* directions. When demand is inelastic, price and total expenditure move in the *same* direction. This relationship is easy to identify in the case of a linear demand curve. Recall that above the midpoint, demand is elastic, and below the midpoint, demand is inelastic. In the upper section of Exhibit 4, total expenditure ($P \times Q$) is measured as the area of a rectangle whose base is Q and height is P. Notice that as price falls, the areas of the inscribed rectangles (each outlined with their own dotted or dashed line) at first grow in size, become largest at the midpoint of the demand curve, and thereafter become smaller as price continues to fall and total expenditure declines toward zero. In the lower section of Exhibit 4, total expenditure is shown for each quantity purchased.

Exhibit 4 Elasticity and Total Expenditure



Note: Figure depicts the relationship among changes in price, changes in quantity, and changes in total expenditure. Maximum total expenditure occurs at the unit-elastic point on a linear demand curve (the cross-hatched rectangle).

The relationships just described hold for any demand curve, so it does not matter whether we are dealing with the demand curve of an individual consumer, the demand curve of the market, or the demand curve facing any given seller. For a market, the total expenditure by buyers becomes the total revenue to sellers in that market. It follows, then, that if market demand is elastic, a fall in price will result in an increase in total revenue to sellers as a whole, and if demand is inelastic, a fall in price will result in a decrease in total revenue to sellers. If the demand faced by any given seller were inelastic at the current price, that seller could increase revenue by increasing its price. But because demand is negatively sloped, the increase in price would decrease total units sold, which would almost certainly decrease total production cost. If raising price both increases revenue and decreases cost, such a move would always be profit enhancing. Faced with inelastic demand, a one-product seller would always be inclined to raise the price until the point at which demand becomes elastic.

2.3 Income Elasticity of Demand

Elasticity is a measure of how sensitive one variable is to change in the value of another variable. Up to this point, we have focused on price elasticity, but the quantity demanded of a good is also a function of consumer income.

Income elasticity of demand is defined as the percentage change in quantity demanded $\left(\%\Delta Q_x^d\right)$ divided by the percentage change in income ($\%\Delta I$), holding all other things constant, as shown in Equation 7:

$$E_I^d = \frac{\%\Delta Q_x^d}{\%\Delta I} \tag{7}$$

The structure of this expression is identical to the structure of own-price elasticity given in Equation 5. (All elasticity measures that we will examine have the same general structure; the only thing that changes is the independent variable of interest.) For example, if the income elasticity of demand for some good has a value of 0.8, we would interpret that to mean that whenever income rises by 1%, the quantity demanded at each price would rise by 0.8%.

Although own-price elasticity of demand will almost always be negative, *income* elasticity of demand can be negative, positive, or zero. Positive income elasticity means that as income rises, quantity demanded also rises. Negative income elasticity of demand means that when people experience a rise in income, they buy less of these goods, and when their income falls, they buy more of the same good.

Goods with positive income elasticity are called "normal" goods. Goods with negative income elasticity are called "inferior" goods. Typical examples of inferior goods are rice, potatoes, or less expensive cuts of meat. We will discuss the concepts of normal and inferior goods in a later section.

In our discussion of the demand curve, we held all other things constant, including consumer income, to plot the relationship between price and quantity demanded. If income were to change, the entire demand curve would shift one way or the other. For normal goods, a rise in income would shift the entire demand curve upward and to the right. For inferior goods, however, a rise in income would result in a downward and leftward shift in the entire demand curve.

2.4 Cross-Price Elasticity of Demand

We previously discussed a good's own-price elasticity. However, the price of another good might also have an impact on the demand for that good or service, and we should be able to define an elasticity with respect to the other price (P_y) as well. That elasticity is called the **cross-price elasticity of demand** and takes on the same structure as own-price elasticity and income elasticity of demand, as represented in Equation 8:

$$E_{p_y}^{d} = \frac{\%\Delta Q_x^d}{\%\Delta P_y} \qquad (8)$$

Note how similar this equation is to the equation for own-price elasticity. The only difference is that the subscript on P is now y, where y indicates some other good. This cross-price elasticity of demand measures how sensitive the demand for good X is to changes in the price of some other good, Y, holding all other things constant. For some pairs of goods, X and Y, when the price of Y rises, more of good X is demanded; the cross-price elasticity of demand is positive. Those goods are referred to as **substitutes**. In economics, if the cross-price elasticity of two goods is positive, they are substitutes, irrespective of whether someone would consider them "similar."

This concept is intuitive if you think about two goods that are seen to be close substitutes, perhaps like two brands of beer. When the price of one of your favorite brands of beer rises, you would probably buy less of that brand and more of a cheaper brand, so the cross-price elasticity of demand would be positive. For substitute goods, an increase in the price of one good would shift the demand curve for the other good upward and to the right.

Alternatively, two goods whose cross-price elasticity of demand is negative are said to be **complements**. Typically, these goods tend to be consumed together as a pair, such as gasoline and automobiles or houses and furniture. When automobile prices fall, we might expect the quantity of autos demanded to rise, and thus we might expect to see a rise in the demand for gasoline.

Whether two goods are substitutes or complements might not be immediately intuitive. For example, grocery stores often put things like coffee on sale in the hope that customers will come in for coffee and end up doing their weekly shopping there as well. In that case, coffee and, say, cabbage could very well empirically turn out to be complements even though we would not think that the price of coffee has any relation to sales of cabbage. Regardless of whether someone would see two goods as related in some fashion, if the cross-price elasticity of two goods is negative, they are complements.

Although a conceptual understanding of demand elasticities is helpful in sorting out the qualitative and directional effects among variables, using an empirically estimated demand function can yield insights into the behavior of a market. For illustration, let us return to our hypothetical individual demand function for gasoline in Equation 2, duplicated here for convenience:

$$Q_x^d = 84.5 - 6.39P_x + 0.25I - 2P_y$$

The quantity demanded of a given good $\left(Q_x^d\right)$ is a function of its own price (P_x), consumer income (I), and the price of another good (P_y).

To derive the market demand function, the individual consumers' demand functions are simply added together. If there were 1,000 individuals who represented a market and they all had identical demand functions, the market demand function would be the individual consumer's demand function multiplied by the number of consumers. Using the individual demand function given by Equation 2, the market demand function would be as shown in Equation 9:

$$Q_x^d = 84{,}500 - 6{,}390P_x + 250I - 2{,}000P_y \quad \textbf{(9)}$$

Earlier, when we calculated own-price elasticity of demand, we needed to choose a price at which to calculate the elasticity coefficient. Similarly, we need to choose actual values for the independent variables—P_x, I, and P_y—and insert these values into the "estimated" market demand function to find the quantity demanded. Choosing €1.48 for P_x, €50 (in thousands) for I, and €20 (in thousands) for P_y, we find that the quantity of gasoline demanded is 47,543 liters per month. We now have everything we need to calculate own-price, income, and cross-price elasticities of demand for our market. Those elasticities are expressed in Equations 10, 11, and 12. Each of those expressions has a term denoting the change in quantity divided by the change in each respective variable: own price, $\Delta Q_x/\Delta P_x$; income, $\Delta Q_x/\Delta I$, and cross price, $\Delta Q_x/\Delta P_y$.

As we stated in the discussion of own-price elasticity, when the relationship between two variables is linear, the change in quantity $\left(\Delta Q_x^d\right)$ divided by the change in own price (ΔP_x), income (ΔI), or cross price (ΔP_y) is equal to the slope coefficient on that other variable. The elasticities are calculated by inserting the slope coefficients from Equation 9 into the elasticity formulas.

Own-price elasticity:

$$E_{p_x}^d = \left(\frac{\Delta Q_x^d}{\Delta P_x}\right)\left(\frac{P_x}{Q_x^d}\right) = \left(-6{,}390\right)\left(\frac{1.48}{47{,}542.8}\right) = -0.20 \quad \textbf{(10)}$$

Income elasticity:

$$E_I^d = \left(\frac{\Delta Q_x^d}{\Delta I}\right)\left(\frac{I}{Q_x^d}\right) = \left(250\right)\left(\frac{50}{47{,}542.8}\right) = 0.26 \quad \textbf{(11)}$$

Cross-price elasticity:

$$E_{p_y}^d = \left(\frac{\Delta Q_x^d}{\Delta P_y}\right)\left(\frac{P_y}{Q_x^d}\right) = (-2000)\left(\frac{20}{47{,}542.8}\right) = -0.84 \qquad (12)$$

In our example, at a price of €1.48, the own-price elasticity of demand is –0.20; a 1% increase in the price of gasoline leads to a decrease in quantity demanded of about 0.20% (Equation 10). Because the absolute value of the own-price elasticity is less than one, we characterize demand as being *inelastic* at that price; for example, an increase in price would result in an increase in total expenditure on gasoline by consumers in that market. The income elasticity of demand is 0.26 (Equation 11): A 1% increase in income would result in an increase of 0.26% in the quantity demanded of gasoline. Because that elasticity is positive (but small), we would characterize gasoline as a normal good. The cross-price elasticity of demand between gasoline and automobiles is –0.84 (Equation 12): If the price of automobiles rose by 1%, the demand for gasoline would fall by 0.84%. We would, therefore, characterize gasoline and automobiles as complements because the cross-price elasticity is negative. The magnitude is quite small, however, so we would conclude that the complementary relationship is weak.

EXAMPLE 1

Calculating Elasticities from a Given Demand Function

An individual consumer's monthly demand for downloadable e-books is given by the equation $Q_{eb}^d = 2 - 0.4P_{eb} + 0.0005I + 0.15P_{hb}$, where Q_{eb}^d equals the number of e-books demanded each month, I equals the household monthly income, P_{eb} equals the price of e-books, and P_{hb} equals the price of hardbound books. Assume that the price of e-books is €10.68, household income is €2,300, and the price of hardbound books is €21.40.

1 Determine the value of own-price elasticity of demand for e-books.

2 Determine the income elasticity of demand for e-books.

3 Determine the cross-price elasticity of demand for e-books with respect to the price of hardbound books.

Solution to 1:

The own-price elasticity of demand is given by $\left(\Delta Q_{eb}^d / \Delta P_{eb}\right)\left(P_{eb} / Q_{eb}^d\right)$. Notice from the demand function that $\Delta Q_{eb}^d / \Delta P_{eb} = -0.4$. Inserting the given variable values into the demand function yields $Q_{eb}^d = 2 - (0.4)(10.68) + (.0005)(2300) + (0.15)(21.4) = 2.088$. So at a price of €10.68, the own-price elasticity of demand equals (–0.4)(10.68/2.088) = –2.046, which is elastic because in absolute value the elasticity coefficient is greater than 1.

Solution to 2:

Recall that income elasticity of demand is given by $\left(\Delta Q_{eb}^d / \Delta I\right)\left(I / Q_{eb}^d\right)$. Notice from the demand function that $\Delta Q_{eb}^d / \Delta I = 0.0005$. Inserting the values for I and Q_{eb}^d yields income elasticity of (0.0005)(2,300/2.088) = 0.551, which is positive, so e-books are a normal good.

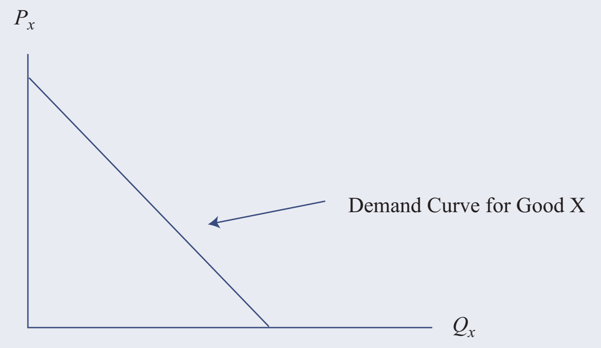
Solution to 3:

Recall that cross-price elasticity of demand is given by $(\Delta Q_{eb}/\Delta P_{hb})(P_{hb}/Q_{eb})$, and notice from the demand function that $\Delta Q_{eb}/\Delta P_{hb} = 0.15$. Inserting the values for P_{hb} and Q_{eb} yields a cross-price elasticity of demand for e-books of (0.15)(21.40/2.088) = 1.537, which is positive, implying that e-books and hardbound books are substitutes.

2.5 Substitution and Income Effects

The law of demand states that if nothing changes other than the price of a particular good or service itself, a decrease in that good's price will tend to result in a greater quantity of that good being purchased. Simply stated, it is the assumption that a demand curve has negative slope; that is, where price per unit is measured on the vertical axis and quantity demanded per time period is measured on the horizontal axis, the demand curve is falling from left to right, as shown in Exhibit 5.

Exhibit 5 A Negatively Sloped Demand Curve—The Law of Demand



There are two reasons why a consumer would be expected to purchase more of a good when its price falls and less of a good when its price rises. These two reasons are known as the substitution effect and the income effect of a change in price. We address these two effects separately and then examine the combination of the two.

When the price of something—say, gasoline—falls, that good becomes relatively less costly compared with other goods or services a consumer might purchase. For example, gasoline is used in driving to work, so when its price falls, it is relatively cheaper to drive to work than to take public transportation. Hence, the consumer is likely to substitute a little more driving to work for a little less public transportation. When the price of beef falls, it becomes relatively cheaper than chicken. The typical consumer is, therefore, likely to purchase a little more beef and a little less chicken.

On its own, the substitution effect suggests that when the price of something falls, consumers tend to purchase more of that good. But another influence is often at work as well—the income effect. Consider a consumer spending all of her "money income" on a given combination of goods and services. (Her money income is simply the quantity of dollars or euros, or other relevant currency, that is available to her to spend in any given time period.) Now suppose the price of something she was regularly purchasing falls while her money income and the prices of all other goods remain unchanged. Economists refer to this as an increase in purchasing power or **real income**. For most goods and services, consumers tend to buy more of them when their income rises. So when the price of a good—say, beef—falls, most consumers would tend to buy more beef because of the increase in their real income. Although the consumer's money income (the number on her paycheck) is assumed not to have changed, her real income has risen because she can now buy more beef—and other goods, too—as a result of the fall in the price of that one good. So, quite apart from the substitution effect of a fall in a good's price, the income effect tends to cause consumers to purchase more of that good as well.

Substitution and income effects work the other way, too. If the price of beef were to rise, the substitution effect would cause the consumer to buy less of it and substitute more chicken for the now relatively more expensive beef. Additionally, the rise in the price of beef results in a decrease in the consumer's real income because now she can buy less goods with the same amount of money income. If beef is a good that consumers tend to buy more of when their income rises and less of when their income falls, then the rise in beef price would have an income effect that causes the consumer to buy less of it.

2.6 Normal and Inferior Goods

Economists classify goods on various dimensions, one of which relates to how consumers' purchases of a good respond to changes in consumer income. Earlier, when discussing income elasticity of demand, we introduced the concept of normal goods and inferior goods. For most goods and services, an increase in income would cause consumers to buy more; these are called **normal goods**. But that does not hold true for all goods: There are goods that consumers buy less of when their income rises and goods that they buy more of when their incomes fall. These are called **inferior goods**. This section will distinguish between normal goods and inferior goods.

We previously discussed income and substitution effects of a change in price. If a good is normal, a decrease in price will result in the consumer buying more of that good. Both the substitution effect and the income effect are at play here:

- A decrease in price tends to cause consumers to buy more of this good in place of other goods—the substitution effect.
- The increase in real income resulting from the decline in this good's price causes people to buy even more of this good when its price falls—the income effect.

So, we can say that for normal goods (restaurant meals, for example, as most people tend to eat out more often when their incomes rise), the substitution and the income effects reinforce one another to cause the demand curve to be negatively sloped.

For inferior goods (cheaper cuts of meat or generic beverages, for example, which most people buy less of as their incomes rise), an increase in income causes consumers to buy less, not more, and if their incomes fall, they buy more, not less. "Inferior" does not imply anything at all about the quality of the good; it is simply used to refer to a good for which an increase in income causes some people to buy less of it.

The same good could be normal for some consumers while it is inferior for others. Consider a very low-income segment of the population. For those consumers, an increase in their income might very well result in their buying more fast-food meals. They might take some of that added income and enjoy eating out at a fast-food restaurant a little more often. Now consider a high-income group. If their income rises, they might be much less inclined to eat at fast-food restaurants and instead do their dining out at a fashionable French bistro, for example. So, fast-food meals might be a normal good for some people and an inferior good for others.

Let us now consider the substitution and income effects of a change in the price of normal and inferior goods. The substitution effect says that if the price of a good falls, the consumer will substitute more of this good in the consumption bundle and buy less of some other good. The substitution effect is true for both normal and inferior goods. Next, we provide an example.

We begin with a hypothetical consumer with a certain money income (R$200,000). Given the prices for all goods, he makes a decision to buy a given amount of Good *X*, coffee. If the price of coffee falls, the consumer is better off than when the price was higher. We can assume that this consumer would have been willing to pay some amount of money each month to be able to buy coffee at the lower price. We now have two states of the world: In State 1, he spends his income on all the various goods, including his desired quantity of coffee at the original price. In State 2, he is able to buy coffee at the new lower price, but because he has paid a portion of his income to buy coffee at the lower price, he now has less money income to spend on all goods combined. If we adjusted the amount of money he would have to pay to lock in the lower price of coffee until he is just indifferent between the two states of the world, we would have exactly offset the "good" thing of the lower price with the "bad" thing of less income. This removes the income effect of the price decrease and allows us to isolate the pure substitution effect. We find that in State 2, he would buy more coffee than in State 1. The pure substitution effect is always in the direction of buying more at the lower relative price.

Continuing our example, assume that we give back to the consumer the amount of money he is willing to pay for the privilege of buying coffee at the lower price. Clearly, he is better off because now he can buy coffee at the lower price without having to pay for the privilege. We want to know whether, with this higher money income, he will now buy more or less coffee at the lower price. The answer depends on whether coffee is a normal or an inferior good for this consumer. Recall that for normal goods, an increase in income causes consumers to buy more, but for inferior goods, an increase in income causes consumers to buy less.

In conclusion, the substitution effect of a change in the price of a good will always be in the direction of buying more at a lower price or less at a higher price. The income effect of that same price change, however, depends on whether the good is normal or inferior. If the good is normal, the income effect reinforces the substitution effect, both leading to a negatively sloped demand curve. But if the good is inferior, the income effect and the substitution effect work in opposite directions; the income effect tends to mitigate the substitution effect.

Exhibit 6 summarizes the substitution and income effects for normal and inferior goods.

Exhibit 6 The Substitution and Income Effects of a Price Decrease on Normal and Inferior Goods

	Substitution Effect	Income Effect
Normal good	Buy more because the good is relatively cheaper than its substitutes.	Buy more because the increase in purchasing power raises the total consumption level.
Inferior good	Buy more because the good is relatively cheaper than its substitutes.	Buy less because the increase in real income prompts the consumer to buy less of the inferior good in favor of its preferred substitutes.

Exceptions to the Law of Demand

In virtually every case in the real world, the law of demand holds: A decrease in price results in an increase in quantity demanded, resulting in a negatively sloped demand curve. In a few unusual cases, however, we may find a positively sloped demand curve—a decrease (increase) in price may result in a decrease (increase) in the quantity demanded. These unusual cases are called Giffen goods and Veblen goods.

In theory, it is possible for the income effect to be so strong and so negative as to overpower the substitution effect. In such a case, more of a good would be consumed as the price rises and less would be consumed as the price falls. These goods are called **Giffen goods**, named for Robert Giffen based on his observations of the purchasing habits of the Victorian era poor. For many decades, no one really believed that a Giffen good actually existed anywhere other than in textbooks. But in recent years, studies have documented a few rare cases. One study was conducted in a poor rural community where individuals spend a very large portion of their incomes on rice. For these individuals, rice was an inferior good. Under the law of demand, the quantity of rice purchased would rise with the decline in price, but the rise in quantity would be partially offset by the income effect (a decrease in the amount of rice purchased as a result of rising incomes). What the experimenters discovered, however, was that for a certain subset of consumers, the quantity of rice purchased declined in absolute terms—the income effect actually overwhelmed the substitution effect. For consumers living at subsistence levels—incomes just barely sufficient to enable them to meet their caloric intake needs—a decline in the price of the staple enabled them to shift more of their consumption from rice to the alternate sources of calories in their diet (e.g., meat).

With some goods, the item's price tag itself might drive the consumer's preferences for it. Thorstein Veblen posited just such a circumstance in his concept of conspicuous consumption. According to this way of thinking, a consumer might derive utility out of being known by others to consume a so-called high-status good, such as a luxury automobile or a very expensive piece of jewelry. Importantly, it is the high price itself that partly imparts value to such a good. These are called **Veblen goods**, and they derive their value from the consumption of them as symbols of the purchaser's high status in society; they are certainly not inferior goods. It is argued that by increasing the price of a Veblen good, the consumer would be more inclined to purchase it, not less.

EXAMPLE 2

Income and Substitution Effects of a Decrease in Price

Monica has a monthly entertainment budget that she spends on (a) movies and (b) an assortment of other entertainment items. When the price of each movie is $8, she spends a quarter of her budget on six movies a month and the rest of her budget on other entertainment. Monica was offered an opportunity to join a movie club at her local theater that allows her to purchase movies at half the regular price, and she can choose each month whether to join the movie club or not. There is a membership fee she must pay for each month she belongs to the club. Monica is exactly indifferent between (a) not buying the membership and, therefore, paying $8 for movies and (b) buying the membership and paying $4 per movie. So, she flips a coin each month to determine whether to join the club that month. In months that she does join the club, she sees eight movies. For her birthday, a friend gave her a one-month club membership as a gift, and that month she saw 12 movies.

1 If there were no club and the price of movies were to simply fall from \$8 to \$4, how many more movies would Monica buy each month?

2 Determine how much Monica is willing to pay each month for the privilege of buying movies at half price. (What is the value of X that makes her indifferent between joining the club and not joining it?)

3 Of the increased number of movies Monica would purchase if the price were to fall from \$8 to \$4, determine how much of the increase would be attributable to the substitution effect and how much to the income effect of that price decrease.

4 For Monica, are movies a normal, inferior, or Giffen good?

Solution to 1:

Six movies. When her friend gave her a club membership, she bought 12 movies instead of her usual 6. With the gift of the club membership, Monica could buy movies at a price of \$4 without paying for that privilege. This is the same as if the price of each movie fell from \$8 to \$4.

Solution to 2:

Note that Monica is indifferent between two states of the world: State A, in which she has all of her entertainment budget to spend on movies and other entertainment but must pay full price of \$8 per movie, and State B, in which she has to pay some dollar amount X for the privilege of buying movies at half price. So, X is the maximum she would pay for a membership fee. She buys eight movies in months when she joins the club. Without a club membership, those movies would cost her \$64 (8 movies × \$8). With a club membership, the movies would cost her \$32 (8 movies × \$4). So the most she is willing to pay for a club membership is \$32. (Note that one might be tempted to say she would be willing to pay only \$24 for the membership because she was buying six movies at \$8, spending \$48, whereas if she were able to buy six movies at only \$4 per movie, she would have to spend only \$24. But because of the substitution effect, she would now be willing to buy more movies than before, so her benefit from the half-price privilege is worth more than \$24.)

Solution to 3:

When Monica pays the club membership herself, she buys eight movies, two more than usual. Because Monica is equally well off whether she joins the club for a monthly fee and thereby pays half price or whether she does not join the club and pays full price, we can say that the income effect of the price decrease has been removed by charging her the monthly fee. So the increase from six movies to eight is the result of the substitution effect. When Monica's friend gave her the gift of a club membership, allowing her to pay half price without paying for the privilege, Monica bought 12 movies, 6 more than usual and 4 more than she would have had she paid the membership fee. The increase from 8 movies to 12 is the result of the income effect.

Solution to 4:

When the price fell from $8 to $4, Monica bought more movies, so clearly movies are not a Giffen good for her. Additionally, because the substitution effect and the income effect are in the same direction of buying more movies, they are a normal good for Monica. The substitution effect caused her to buy two more movies, and the income effect caused her to buy an additional four movies.

SUPPLY ANALYSIS: THE FIRM

3

To fully comprehend the supply side of a consumer goods market, an analyst must understand the firm's costs. (As a reminder, this reading builds on the basics of the market model as covered in the recommended prerequisite reading material.)

The firm's marginal cost is the foundation of the firm's ability and willingness to offer a given quantity for sale, and its costs depend on both the productivity of its inputs and their prices. In this section, we will describe the firm's cost curves—total, average, and marginal costs in both the short run and in the long run—paying special attention to what economists call the **law of diminishing marginal returns**. We will then use this information to explore the conditions under which a firm would find it beneficial to continue operation, even if its economic profits are negative, and at what levels of production its shutdown and breakeven points occur. Long-run costs will be examined in the context of economies and diseconomies of scale.

3.1 Marginal Returns and Productivity

There is an economic phenomenon known as **increasing marginal returns**, in which **marginal product**—the productivity of each additional unit of a resource—increases as additional units of that input are employed.

Initially, a firm can experience increasing returns from adding labor to the production process because of the concepts of specialization and division of labor. At first, by having too few workers relative to total physical capital, the understaffing situation requires employees to multi-task and share duties. As more workers are added, employees can specialize, become more adept at their individual functions, and realize an increase in marginal productivity. But after a certain output level, the law of diminishing marginal returns becomes evident.

When more and more workers are added to a fixed capital base, the marginal return of the labor factor eventually decreases because the fixed input restricts the output potential of additional workers. As an illustration, consider automobile production. When an auto manufacturing plant is operating at full capacity, adding additional labor will not increase production because the physical plant is already 100% employed. More labor hours will merely add to costs without adding to output. Assuming all workers are of equal quality and motivation, the decline in marginal product occurs in the short run, where all other resources (typically, plant size, physical capital, and technology) are fixed.

Marginal returns are directly related to **input productivity**, a measure of the output per unit of input.

3.1.1 *Productivity: The Relationship between Production and Cost*

The cost of producing anything depends on the amount of *inputs*, or *factors of production* (these terms are synonymous), and the input prices. Examples of factors of production are employee hours, machine hours, raw materials, and so on. For simplicity, economists typically concentrate on only two inputs, labor and capital, although obviously there can be many inputs to a particular production process. The labor input is simply employee time, and it is measured as labor hours per time period, such as per week or per month. We denote labor hours as L. If a firm is using two laborers per week and each laborer works 35 hours per week, then L equals 70 labor hours per week. We denote hours of capital as K. If the firm is using three machines and each one is used for 12 hours per week, then K equals 36 machine hours per week. That is, the capital input is measured as machine hours used per time period. In this way, capital and labor are stated in similar terms. They represent flows of services—labor hours and machine hours—that are used to produce a flow of output per time period.

Accordingly, the respective input prices would be the wage rate per labor hour (we use w to denote wage rate) and the rental rate per machine hour (we use r to denote the rental rate per machine hour). It is helpful to think of a firm as renting the services of labor and of machines. Although the firm might own its own machines, it could in theory rent its machines out to another user, so it is forgoing the rate it could earn elsewhere when it is using its machines internally instead of renting them out. So, a firm is not using its own machines "for free." It is incurring the **opportunity cost** of not being able to rent those machines to another user.

The **total cost** of production (TC) is the number of hours of labor multiplied by the wage rate plus the number of machine hours multiplied by the rental rate of machines:

$$\text{TC} = (w)(L) + (r)(K)$$

This formula illustrates that the total cost is just the cost of all the firm's inputs. It is not a cost function, however, which is a relationship between the cost of production and the flow of output. The cost function $C = \text{f}(Q)$, where (Q) denotes the flow of output in units of production per time period, relates the production cost per time period to the number of units of output produced per time period.

Two things could cause the cost of producing any given level of output to fall: Either the price of one or both inputs could fall or the inputs themselves could become more productive and less of them would be needed (e.g., a worker is more productive when fewer hours of labor are needed to produce the same output). The reverse is true also: A rise in cost could result from either a rise in input prices or a fall in input productivity, or both.

Why is productivity important? Cost-minimization and profit-maximization behavior dictate that the firm strives to maximize productivity—for example, produce the most output per unit of input or produce any given level of output with the least amount of inputs. A firm that lags behind the industry in productivity is at a competitive disadvantage and is likely to face decreases in future earnings and shareholders' wealth. An increase in productivity lowers production costs, which leads to greater profitability and investment value. These productivity benefits can be fully or partially distributed to other stakeholders of the business, such as to consumers in the form of lower prices and to employees in the form of enhanced compensation. Transferring some or all of the productivity rewards to non-equity holders creates synergies that benefit shareholders over time.

The benefits from increased productivity are as follows:

- lower business costs, which translate into increased profitability;
- an increase in the market value of equity and shareholders' wealth resulting from an increase in profit; and
- an increase in worker rewards, which motivates further productivity increases from labor.

Undoubtedly, increases in productivity reinforce and strengthen the competitive position of the firm over the long run. A fundamental analysis of a company should examine the firm's commitment to productivity enhancements and the degree to which productivity is integrated into the competitive nature of the industry or market. In some cases, productivity is not only an important promoter of growth in firm value over the long term but is also the key factor for economic survival. A business that lags the market in terms of productivity often finds itself less competitive, while at the same time confronting profit erosion and deterioration in shareholders' wealth. Typical productivity measures for a firm are based on the concepts of total product, average product, and marginal product of labor.

3.1.2 *Total, Average, and Marginal Product of Labor*

When measuring a firm's operating efficiency, it is easier and more practical to use a single resource factor as the input variable rather than a bundle of the different resources that the firm uses in producing units of output. As discussed in the previous section, labor is typically the input that is the most identifiable and calculable for measuring productivity. However, any input that is not difficult to quantify can be used. As an example, a business that manually assembles widgets has 50 workers, one production facility, and an assortment of equipment and hand tools. The firm would like to assess its productivity when using these three input factors to produce widgets. In this example, it is most appropriate to use labor as the input factor for determining productivity because the firm uses only one (fixed) plant building and a variety of other physical capital.

We will use labor as the input variable to illustrate the concepts of total product, average product, and marginal product. Exhibit 7 provides a summary of these three concepts.

Exhibit 7 Definitions and Calculations for Total, Marginal, and Average Product of Labor

Term	Calculation
Total product	Sum of the output from all inputs during a time period; usually illustrated as the total output (Q) using labor quantity (L)
Average product	Total product divided by the quantity of a given input; measured as total product divided by the number of worker hours used at that output level (Q/L)
Marginal product	The amount of additional output resulting from using one more unit of input assuming other inputs are fixed; measured by taking the difference in total product and dividing by the change in the quantity of labor ($\Delta Q/\Delta L$)

Total product (Q) is defined as the aggregate sum of production for a firm during a time period. As a measure of productivity, total product provides superficial information about how effective and efficient a firm is in terms of producing output. For instance, three firms—Company A, Company B, and Company C—that make up an entire industry have total output levels of 100,000 units, 180,000 units, and 200,000 units, respectively. Obviously, Company C dominates the market with a 41.7% share, followed by Company B's 37.5% share and Company A's 20.8% portion of the market. However, this information says little about how efficient each firm is in generating its total output level. Total product only provides insight into a firm's production volume relative to the industry; it does not show how efficient a firm is in producing its output.

Average product of labor (AP_L) measures the productivity of an input (in this case, labor) on average and is calculated by dividing total product by the total number of units for the given input that is used to generate that output. Average product is usually measured on the basis of the labor input. It is a representative or overall measure of labor's productivity: Some workers are more productive than average, and others are less productive than average.

Exhibit 8 compares the productivity of the three firms introduced earlier. Company A employs 100 worker hours and produces 100,000 widgets per hour. Company B employs 200 worker hours and produces 180,000 widgets per hour. Company C employs 250 worker hours and produces 200,000 widgets per hour.

Exhibit 8 Comparing Productivity

	Output (Q)	Number of Worker Hours (L)	Average Product of Labor (AP_L)
Company A	100,000	100	1,000
Company B	180,000	200	900
Company C	200,000	250	800

Using this metric, it is apparent that Company A, with AP_L equal to 1,000, is the most efficient firm, despite having the lowest market share. Company C has the largest market share, but it is the least efficient of the three, with AP_L equal to 800. Assuming that Company A can maintain its productivity advantage over the long run, it will be positioned to generate the greatest return on investment through lower costs and higher profit outcomes relative to the other firms in the market.

Marginal product of labor (MP_L), also known as *marginal return*, measures the productivity of each additional unit of input and is calculated by observing the difference in total product when adding another unit of input (assuming other resource quantities are held constant). It is a gauge of the productivity of the individual additional worker hour rather than an average across all workers.

Exhibit 9 provides a numerical illustration for total, average, and marginal products of labor.

Exhibit 9 Total, Average, and Marginal Product of Labor

Labor (L)	Total Product (Q_L)	Average Product (AP_L)	Marginal Product (MP_L)
0	0	—	—
1	100	100	100
2	210	105	110
3	300	100	90
4	360	90	60
5	400	80	40
6	420	70	20
7	350	50	–70

Total product increases as the firm adds each additional hour of labor—until the seventh labor hour, at which point total production declines by 70 units. Obviously, the firm would want to avoid negative worker productivity.

At an employment level of five labor hours, AP_L is 80 units (400/5) and MP_L is 40 units [(400 – 360)/(5 – 4)]. The average productivity for all five labor hours is 80 units, but the productivity of the fifth labor hour is only 40 units.

EXAMPLE 3

Calculation and Interpretation of Total, Average, and Marginal Product

Exhibit 10 illustrates the production relationship between the number of machine hours and total product.

1 Interpret the results for total, average, and marginal product.

2 Indicate at what point increasing marginal returns change to diminishing marginal returns.

Exhibit 10

Machine Hours (K)	Total Product (Q_K)	Average Product (AP_K)	Marginal Product (MP_K)
0	0	—	—
1	1,000	1,000	1,000
2	2,500	1,250	1,500
3	4,500	1,500	2,000
4	6,400	1,600	1,900
5	7,400	1,480	1,000
6	7,500	1,250	100
7	7,000	1,000	–500

Solution to 1:

Total product increases up to six machine hours, where it tops out at 7,500. Because total product declines from Hour 6 to Hour 7, the marginal product for Machine Hour 7 is negative 500 units. Average product peaks at 1,600 units with four machine hours. Average product increases at a steady pace with the addition of Machine Hours 2 and 3. The addition of Machine Hour 4 continues to increase average product but at a decreasing rate. Beyond four machine hours, average product decreases—at an increasing rate. Marginal product peaks with Machine Hour 3 and decreases thereafter.

Solution to 2:

The marginal product, MP_K, of Machine Hour 3 is 2,000. The marginal product of each additional machine hour beyond Machine Hour 3 declines. Diminishing marginal returns are evident beyond Machine Hour 3.

A firm has a choice of using total product, average product, marginal product, or some combination of the three to measure productivity. Because total product is simply an indication of a firm's output volume and potential market share, average product and marginal product are better gauges of a firm's productivity. Both can reveal competitive advantage through production efficiency. However, individual worker productivity is not easily measurable when workers perform tasks collectively. In this case, average product is the preferred measure of productivity performance.

Referring to the total product column in Exhibit 9, output is more than twice as great (210 widgets) when two hours of labor are used as opposed to only one hour (100 widgets.) In this range of production, there is an increase in return when employee hours are added to the production process. This is the phenomenon of increasing marginal returns.

3.2 Breakeven and Shutdown Analysis

Two important considerations of any firm are its level of profitability and whether to continue to operate in the current environment. Economists define profit differently than do accountants. **Economic profit** is defined as the difference between total revenue (TR) and total **economic costs**. **Accounting profit** is the difference between TR and total **accounting cost**. TR is the same from both an accounting standpoint and an economic standpoint; it is derived by multiplying the selling price per unit of output by the number of units: TR = $(P)(Q)$. The difference between the two measures of profit, therefore, lies in an understanding of economic cost (also called "opportunity cost," which is defined in detail in the next section).

3.2.1 *Economic Cost vs. Accounting Cost*

The opportunity cost of any particular decision, such as to produce a given level of output, can be determined by measuring the benefit forgone by not implementing the next best alternative. Suppose that a firm is currently operating with hired labor and its own plant and equipment to produce output at some level. The firm must continuously decide either to keep the level of output the same or to change it. The decision to maintain the same output requires that the firm hire the same amount of labor input and use the same level of its capital inputs as before. The labor expense is both an economic cost and an accounting cost because the money spent on labor hours could have been used for something else (opportunity cost), and it is also a current expense for the firm (accounting cost.)

Accountants typically attempt to recognize the cost of plant and equipment in the form of accounting depreciation, which is a means of distributing the historical cost of the fixed capital among the units of production for financial reporting purposes. The money spent in the past on the firm's plant and equipment is what economists call a "sunk cost." Because sunk costs cannot be altered, they cannot affect an optimal decision, which is forward looking. Sunk costs are therefore ignored, and the key management question is, Going forward, what are the opportunity costs and benefits of maintaining a given level of output?

Here is where economic depreciation comes into play. To understand the opportunity costs of using our plant and equipment—already bought and paid for—for one more period of time to produce output, we have to ask the question, What else could be done with that fixed capital if it were not used to produce our output? The answer might be that because there is no external market for our machines and buildings, we are forgoing nothing by using it to produce output. Or it might be that there is a market where we could rent out or sell our capital equipment elsewhere instead of using it to produce output. That rental rate is the economic depreciation associated with using our own equipment to produce output instead of renting or selling it elsewhere.

Economic depreciation is forward looking. It asks, What am I giving up if I use my resources to produce output in the coming period? Accounting depreciation is backward looking. It asks, How should I distribute the historical cost—that I have already paid—across units of output that I intend to produce this period? Both concepts are useful—one for making managerial decisions about output and the other as a way spreading historical costs for reporting or tax purposes—but there is not necessarily a direct relationship between the two.

3.2.2 *Marginal Revenue, Marginal Cost, and Profit Maximization*

It is assumed that any for-profit firm's management is tasked with achieving the goal of shareholder wealth maximization. Put most simply, that translates into the goal of economic profit maximization. Hereafter, when the word *profit* is used, it will be economic profit that we have in mind. Because profit is defined as TR minus TC, anything that increases revenue more than cost or decreases cost more than revenue will increase profit. Before we address profit maximization, we must introduce two important concepts: marginal revenue and marginal cost.

Marginal revenue (MR) is the additional revenue the firm realizes from the decision to increase output by one unit per time period. That is, MR = $\Delta TR/\Delta Q$. If the firm is operating in what economists call a perfectly competitive market, it is one of many sellers of identical products in an environment characterized by low or non-existent barriers to entry. Under perfect competition, the firm has no pricing power because there are many perfect substitutes for the product it sells. If it were to attempt to raise the price even by a very small amount, it would lose all of its sales to competitors. On the other hand, it can sell essentially any amount of product it wants without lowering the price below the market price.

Take the wheat market as an example of a perfectly competitive market. A seller of wheat would have no control over the market price of wheat; thus, because TR = $(P)(Q)$, MR for this firm is simply price per unit of output. This firm is said to face a perfectly horizontal (zero-sloped), or infinitely elastic, demand curve for its product. For example, if the firm is selling 1,000 bushels of wheat per week at a price of £3 per bushel, TR is £3,000. If the firm were to increase its output by one unit, then TR would rise by exactly £3 because the firm would not have to lower its price to sell that added unit. So, for sellers in a market with perfect competition, MR = P.

In contrast, if a firm sells a product that is differentiated from other firms' products and that has a large market share, the firm is said to be operating in an environment of imperfect competition. In the extreme case of imperfect competition, there might be only one firm selling a product with no close substitutes. That firm holds a monopoly, and it is subject to the market demand curve for its product. Whether a monopoly or simply operating under imperfect competition, the firm faces a negatively sloped demand curve and must lower its price to sell another unit. Thus, MR will be lower than price.

To illustrate this concept, we will decompose MR. Recall from earlier in the reading that

$$TR = (P)(Q)$$

and

$$MR = \Delta TR/\Delta Q$$

Change in total revenue (ΔTR), the numerator of the ratio, can be written as $(P)(\Delta Q) + (Q)(\Delta P)$.

There are two competing forces affecting revenue: (1) Additional units are sold at the new price, and (2) all units must now be sold at the lower price. The firm is selling more units, but it is selling all units at a lower price than before.

To find MR, we divide the change in TR by the change in quantity:

$$MR = \frac{(P)(\Delta Q)}{\Delta Q} + \frac{(Q)(\Delta P)}{\Delta Q} = P + Q\frac{(\Delta P)}{\Delta Q}$$

In other words, MR is equal to price but with an "adjustment" equal to $(Q)(\Delta P/\Delta Q)$.

Taking this one step further, recall that earlier we said $(\Delta P/\Delta Q)$ is the slope of the demand curve. From our expression just given, MR = $P + Q(\Delta P/\Delta Q)$; so, MR is equal to price with an adjustment equal to quantity times the slope of the demand curve.

A perfectly competitive firm faces a demand curve with a slope of zero. Substituting 0 for $\Delta P/\Delta Q$ into the expression given, it becomes clear that MR is equal to price for the perfectly competitive firm—it need not lower its price to sell an additional unit. For a firm in an imperfectly competitive market, however, the demand curve is negatively sloped ($\Delta P/\Delta Q < 0$). Substituting this negative number into the expression for MR, $P + Q(\Delta P/\Delta Q)$, it becomes clear that MR for an imperfectly competitive firm is less than price.

Marginal cost (MC) is the increase to total cost resulting from the firm's decision to increase output by one additional unit per time period: $\text{MC} = \Delta\text{TC}/\Delta Q$. Economists distinguish between short-run marginal cost (SMC) and long-run marginal cost (LMC). Labor is variable over the short run, but the quantity of capital cannot be changed in the short run because there is a lead time required to build or buy new plant equipment and put it in place. In the long run, all inputs are variable.

SMC is essentially the additional cost of the variable input, labor, that must be incurred to increase the level of output by one unit. LMC is the additional cost of all inputs necessary to increase the level of output, allowing the firm the flexibility of changing both labor and capital inputs in a way that maximizes efficiency.

Understanding MC is aided by recalling that cost is *directly* related to input prices and *inversely* related to productivity. For example, if the wage rate were to rise, cost would also rise. If labor were to become more productive, cost would fall. This relationship can be captured in an expression that relates SMC to wage rate (w) and MP_L: $\text{SMC} = w/\text{MP}_L$.

This relationship between cost and productivity also holds with average variable cost. **Variable costs** are all costs that fluctuate with the level of production and sales. **Average variable cost** (AVC) is the ratio of total variable cost to total output: $\text{AVC} = \text{TVC}/Q$. Again, if labor's wage rises, AVC also rises; but if labor were to become more productive, AVC falls. This relationship is captured by the expression $\text{AVC} = w/\text{AP}_L$.

Earlier, we noted that over some range of low output, the firm might benefit from increasing marginal productivity of its labor input as workers begin to specialize. As the MP_L increases, SMCs decline. Eventually, as more and more labor is added to a fixed amount of capital, the MP_L must fall, causing SMCs to rise.

We began this section by stating that the goal of management is to maximize profit. We now address the conditions necessary for reaching that goal. Consider a firm currently producing 1,000 widgets each week and whose management is contemplating increasing that output incrementally. Would that additional unit increase profit? Clearly, profit would be increased (or losses reduced) if the additional revenue from that next unit were greater than the additional cost. So, a profit-seeking firm should increase Q if $\text{MR} > \text{MC}$. Conversely, if the additional unit added more to cost than to revenue, the firm should reduce output because it would save more in cost than it would lose in revenue. Only if the additional cost were exactly equal to the additional revenue would the firm be maximizing its profit.

There is another condition (called a second-order condition) necessary for profit maximization: At the level of output at which MR = MC, MC cannot be falling. This condition is fairly intuitive. If MC is falling with additional output, MP_L would be rising. (Recall that $\text{SMC} = w/\text{MP}_L$) If one additional hour of labor input causes MC to fall, the firm would want to add that hour and continue adding labor until SMC becomes positively sloped. We can sum up the profit-maximization decision for an operating firm as follows: Produce the level of output such that (1) MR = MC and (2) MC is not falling.

3.2.3 *Understanding the Interaction between Total, Variable, Fixed, and Marginal Cost and Output*

Exhibit 11 shows the graphical relationships between total cost, total fixed cost, and total variable cost. TC is the summation of all costs, where costs are classified on the basis of whether they are fixed or variable. **Total fixed cost** (TFC) is the summation of all expenses that do not change as the level of production varies. **Total variable cost** (TVC) is the summation of all variable expenses; TVC rises with increased production and falls with decreased production. At zero production, TC is equal to TFC because TVC at this output level is zero. The curve for TC always lies parallel to and above the TVC curve by the amount of TFC.

Exhibit 11 Total Cost, Total Variable Cost, and Total Fixed Cost

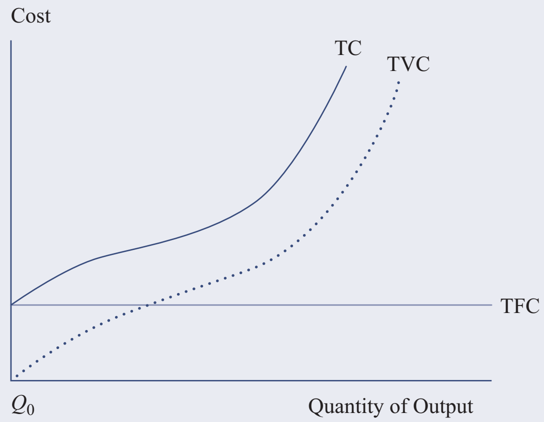
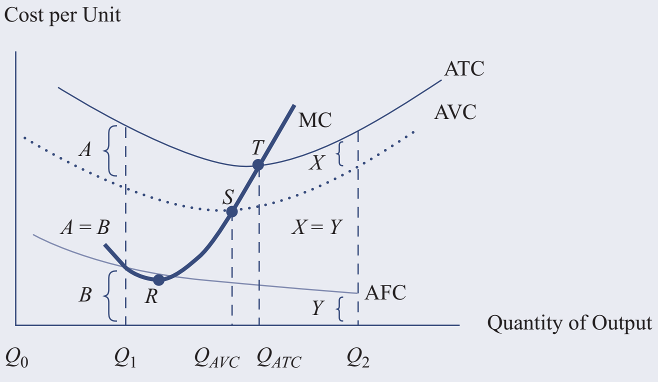


Exhibit 12 shows the relationships between the **average total cost** (ATC), average variable cost (AVC), **average fixed cost** (AFC), and marginal cost (MC) curves in the short run. As output quantity increases, AFC declines because TFCs are spread over a larger number of units. Both ATC and AVC take on a bowl-shaped pattern in which each curve initially declines, reaches a minimum average cost output level, and then increases after that point. The MC curve intersects both the ATC and the AVC at their minimum points—points *S* and *T*. When MC is less than AVC, AVC will be decreasing. When MC is greater than AVC, AVC will be increasing.

Exhibit 12 Average Total Cost, Average Variable Cost, Average Fixed Cost, and Marginal Cost



S, the lowest point on the AVC curve, is where MC equals AVC. Beyond quantity Q_{AVC}, MC is greater than AVC; thus, the AVC curve begins to rise. Note that it occurs at a quantity lower than the minimum point on the ATC curve.
T, the lowest point on the ATC curve, is where MC equals ATC. Beyond quantity Q_{ATC}, MC is greater than ATC; thus, the ATC curve is rising.
A, the difference between ATC and AVC at output quantity Q_1, is the amount of AFC.
R indicates the lowest point on the MC curve. Beyond this point of production, fixed input constraints reduce the productivity of labor.
X indicates the difference between ATC and AVC at quantity Q_2. It is less than *A* because AFC (*Y*) falls with output.

Exhibit 13 shows an example of how total, average, and marginal costs are derived. TC is calculated by summing TFC and TVC. MC is derived by observing the change in TC as the quantity variable changes. There is a relationship that always holds for average and marginal costs: If MC is less than average cost, average cost must fall, and if MC is greater than average cost, average cost must rise. For example, in Exhibit 13, AVC begins to increase as output rises from 2 to 3 units because MC (50) is greater than AVC (41.7). Also from Exhibit 13, ATC declines up to 3 units because MC is less than ATC. After 3 units, ATC increases because the MC of Unit 4 (85) exceeds the ATC of all prior units (75). Initially, the MC curve declines because of increasing marginal returns to labor, but at some point, it begins to increase because of the law of diminishing marginal returns.

Exhibit 13 Total, Average, Marginal, Fixed, and Variable Costs

Quantity (*Q*)	TFC[a]	AFC	TVC	AVC	TC	ATC	MC
0	100	—	0	—	100	—	—
1	100	100.0	50	50.0	150	150.0	50
2	100	50.0	75	37.5	175	87.5	25
3	100	33.3	125	41.7	225	75.0	50
4	100	25.0	210	52.5	310	77.5	85
5	100	20.0	300	60.0	400	80.0	90
6	100	16.7	450	75.0	550	91.7	150
7	100	14.3	650	92.9	750	107.1	200
8	100	12.5	900	112.5	1,000	125.0	250

Exhibit 13 (Continued)

Quantity (*Q*)	TFC[a]	AFC	TVC	AVC	TC	ATC	MC
9	100	11.1	1,200	133.3	1,300	144.4	300
10	100	10.0	1,550	155.0	1,650	165.0	350

[a] Includes all opportunity costs.

As stated earlier, TC increases as the firm expands output and decreases when production is cut. TC increases at a decreasing rate up to a certain output level. Thereafter, the rate of increase accelerates as the firm gets closer to full utilization of capacity. The rate of change in TC mirrors the rate of change in TVC. In Exhibit 13, TC at 5 units is 400—of which 300 is variable cost and 100 is fixed cost. At 10 units, TC is 1,650—of which 1,550 is variable cost and 100 is fixed cost.

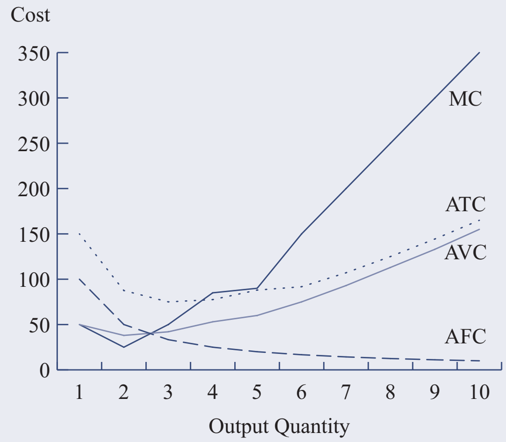
Fixed costs typically are incurred whether the firm produces anything or not. Fixed costs may stay the same over a given range of production but can change to another constant level when production moves outside of that range. The latter is referred to as a **quasi-fixed cost**, although it remains categorized as part of TFC. Examples of fixed costs are debt service, real estate lease agreements, and rental contracts. **Normal profit** is also considered to be a fixed cost because it is a return required by investors on their equity capital regardless of output level. Quasi-fixed cost examples would be certain utilities and administrative salaries that could be lower or avoided altogether when output is zero but would rise to higher constant levels over different production ranges.

Other fixed costs evolve primarily from investments in such fixed assets as real estate, production facilities, and equipment. These fixed costs cannot be arbitrarily cut when production declines. When a firm downsizes, the last expense to be cut is usually fixed cost.

TVC has a direct relationship with quantity. When quantity increases, TVC increases; when quantity decreases, TVC declines. At zero production, TVC is always zero. Variable cost examples are payments for labor, raw materials, and supplies. The change in TVC declines up to a certain output point and then increases as production approaches capacity limits. In Exhibit 13, TVC increases with an increase in quantity. However, the change from 1 to 2 units is 25 (75 – 50), and the change from 9 to 10 units is 350.

Exhibit 14 illustrates the relationships between MC, ATC, AVC, and AFC for the data presented in Exhibit 13.

Exhibit 14 Average Total Cost, Average Variable Cost, Average Fixed Cost, and Marginal Cost for Exhibit 13 Data



Dividing TFC by quantity yields AFC. AFC decreases throughout the production span, reflecting the spreading of a constant cost over more and more production units. At high production volumes, AFC may be so low that it is a small proportion of ATC. In Exhibit 13, AFC declines from 100 at 1 unit to 20 at 5 units, and then to 10 at an output level of 10 units.

In Exhibit 13, AVC at 5 units is 60 (300/5). Over an initial range of production, AVC declines and then reaches a minimum point. Thereafter, AVC increases as the firm uses more of its production capacity. This higher cost results primarily from production constraints imposed by the fixed assets at higher volume levels. The lowest AVC quantity does not correspond to the least-cost quantity for ATC because AFC is still declining. In Exhibit 13, AVC is minimized at 2 units, whereas ATC is minimized at 3 units.

ATC is calculated by dividing TC by quantity (or by summing AFC and AVC). In Exhibit 13, at 3 units, ATC is 75 (TC of 225/3 units of production or AFC of 33.3 + AVC of 41.7). This is the least average cost point of production and the minimum point on the ATC curve. Although cost-minimizing behavior on the part of the firm would dictate operating at the minimum point on its ATC curve, the profit-maximizing quantity may not correspond to this minimum ATC point. Profit per unit, but not necessarily total profit, is maximized at this point.

EXAMPLE 4

Calculation and Interpretation of Total, Average, Marginal, Fixed, and Variable Costs

The first three columns of Exhibit 15 display data on quantity, TFC, and TVC, which are used to calculate TC, AFC, AVC, ATC, and MC. Examine the results for total, average, marginal, fixed, and variable costs. Identify the quantity levels at which the ATC, AVC, and MC values reach their minimum points. Explain the relationship between TFC and TC at a quantity of zero output.

Exhibit 15

Q	TFC[a]	TVC	AFC	AVC	TC	ATC	MC
0	5,000	0	—	—	5,000	—	—
1	5,000	2,000	5,000.0	2,000	7,000	7,000.0	2,000
2	5,000	3,800	2,500.0	1,900	8,800	4,400.0	1,800
3	5,000	5,400	1,666.7	1,800	10,400	3,466.7	1,600
4	5,000	8,000	1,250.0	2,000	13,000	3,250.0	2,600
5	5,000	11,000	1,000.0	2,200	16,000	3,200.0	3,000
6	5,000	15,000	833.3	2,500	20,000	3,333.3	4,000
7	5,000	21,000	714.3	3,000	26,000	3,714.3	6,000
8	5,000	28,800	625.0	3,600	33,800	4,225.0	7,800
9	5,000	38,700	555.6	4,300	43,700	4,855.6	9,900
10	5,000	51,000	500.0	5,100	56,000	5,600.0	12,300

[a] Includes all opportunity costs

Solution:

TFC remains unchanged at 5,000 throughout the entire production range, whereas AFC continuously declines from 5,000 at 1 unit to 500 at 10 units. Both AVC and MC initially decline and then reach their lowest level at 3 units, with costs of 1,800 and 1,600, respectively. Beyond 3 units, both AVC and MC increase, indicating that the cost of production rises with greater output. The least-cost point for ATC is 3,200 at 5 units. At zero output, TC is 5,000, which equals the amount of TFC (at zero output, the firm will need no variable inputs, but it is committed to its fixed plant and equipment in the short run).

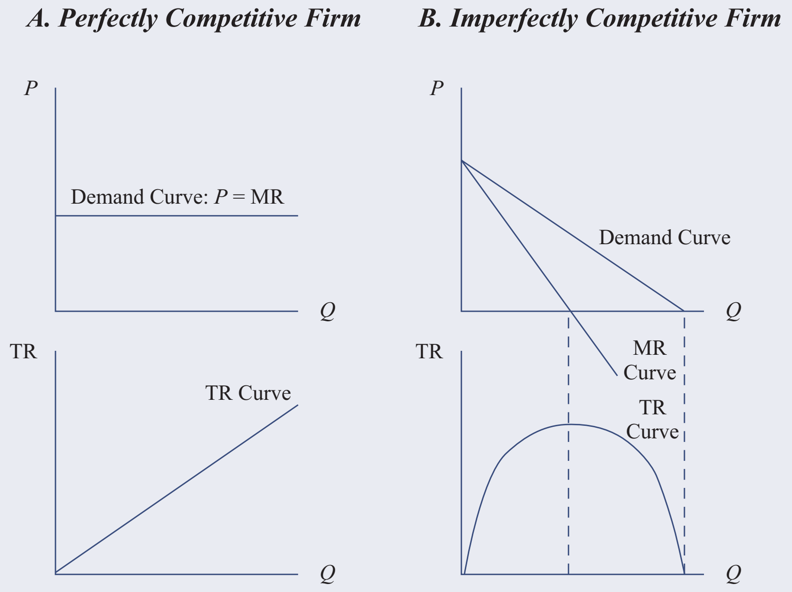
3.2.4 *Revenue under Conditions of Perfect and Imperfect Competition*

Recall from our earlier discussion of profit-maximizing conditions that a firm can generally be classified as operating in either a perfectly competitive or an imperfectly competitive environment. The difference between the two manifests itself in the slope of the demand curve facing the firm. If the environment of the firm is perfectly competitive, it must take the market price of its output as given, so it faces a perfectly elastic, horizontal demand curve. In this case, as we saw previously, the firm's MR and the price of its product are identical. Additionally, the firm's **average revenue** (AR), or revenue per unit, is also equal to price per unit. However, a firm that faces a negatively sloped demand curve must lower its price to sell an additional unit, so its MR is less than price (P).

These characteristics of MR are also applicable to the TR functions. Under conditions of perfect competition, TR (as always) is equal to price times quantity: TR = $(P)(Q)$. But under conditions of perfect competition, price is dictated by the market; the firm has no control over price. As the firm sells one more unit, its TR rises by the exact amount of price per unit.

Under conditions of imperfect competition, price is a variable under the firm's control, and therefore price is a function of quantity: $P = f(Q)$, and TR $= f(Q) \times Q$. For simplicity, suppose the firm is monopolistic and faces the market demand curve, which we will assume is linear and negatively sloped. Because the monopolist is the only seller, its TR is identical to the total expenditure of all buyers in the market. Earlier, we noted what happens as price is reduced and quantity sold increases in this environment: At first, a decrease in price increases total expenditure by buyers and TR to the firm because the decrease in price is outweighed by the increase in units sold. But as price continues to fall, the decrease in price overshadows the increase in quantity, and total expenditure (revenue) falls. We can now depict the demand and TR functions for firms under conditions of perfect and imperfect competition, as shown in Exhibit 16.

Exhibit 16 Demand and Total Revenue Functions for Firms under Conditions of Perfect and Imperfect Competition

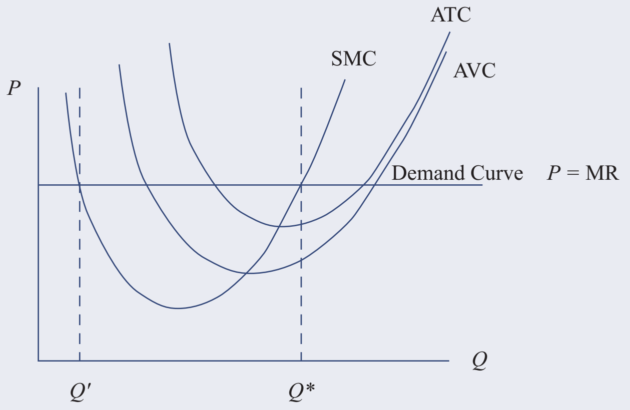


Panel A of Exhibit 16 depicts the demand curve (upper graph) and total revenue curve (lower graph) for the firm under conditions of perfect competition. Notice that the vertical axis in the upper graph is price per unit (e.g., £/bushel), whereas TR is measured on the vertical axis in the lower graph (e.g., £/week.) The same is true for the respective axes in Panel B, which depicts the demand and total revenue curves for the monopolist. The TR curve for the firm under conditions of perfect competition is linear, with a slope equal to price per unit. The TR curve for the monopolist first rises (in the range where MR is positive and demand is elastic) and then falls (in the range where MR is negative and demand is inelastic) with output.

3.2.5 *Profit-Maximization, Breakeven, and Shutdown Points of Production*

We can now combine the firm's short-run TC curves with its TR curves to represent profit maximization in the cases of perfect competition and imperfect competition. Exhibit 17 shows both the AR and average cost curves in one graph for the firm under conditions of perfect competition.

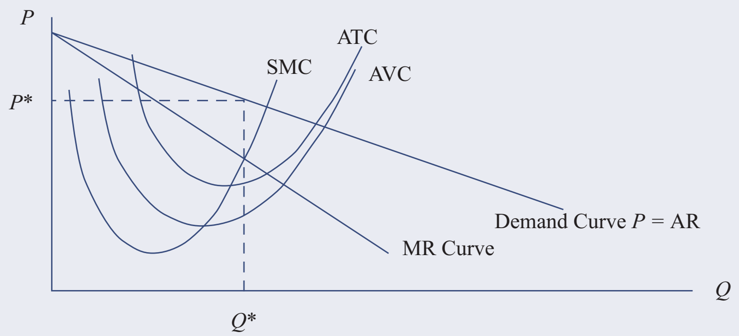
Exhibit 17 Demand and Average and Marginal Cost Curves for the Firm under Conditions of Perfect Competition



The firm is maximizing profit by producing Q^*, where price is equal to SMC and SMC is rising. (Note that there is another output level, Q', where P = SMC, but at that point, SMC is still falling, so this cannot be a profit-maximizing solution.) If market price were to rise, the firm's demand and MR curve would simply shift upward, and the firm would reach a new profit-maximizing output level to the right of Q^*. If, on the other hand, market price were to fall, the firm's demand and MR curve would shift downward, resulting in a new and lower level of profit-maximizing output. As depicted, this firm is currently earning a positive economic profit because market price exceeds ATC at output level Q^*. This profit is possible in the short run, but in the long run, competitors would enter the market to capture some of those profits and would drive the market price down to a level equal to each firm's ATC.

Exhibit 18 depicts the cost and revenue curves for the monopolist that is facing a negatively sloped market demand curve. The MR and demand curves are not identical for this firm. But the profit-maximizing rule is still the same: Find the level of Q that equates SMC to MR—in this case, Q^*. Once that level of output is determined, the optimal price to charge is given by the firm's demand curve at P^*. This monopolist is earning positive economic profit because its price exceeds its ATC. The barriers to entry that give this firm its monopolistic power mean that outside competitors would be unable to compete away this firm's profits.

Exhibit 18 Demand and Average and Marginal Cost Curves for the Monopolistic Firm



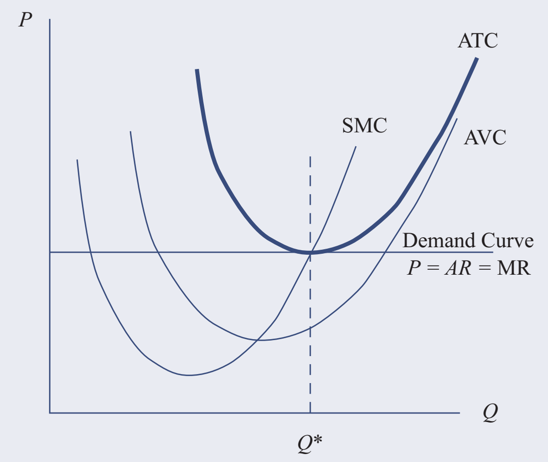
3.2.6 *Breakeven Analysis*

A firm is said to break even if its TR is equal to its TC. It can also be said that a firm breaks even if its price (AR) is exactly equal to its ATC, which is true under conditions of perfect and imperfect competition. Of course, the goal of management is not just to break even but to maximize profit. However, perhaps the best the firm can do is cover all of its economic costs. Economic costs are the sum of total accounting costs and implicit opportunity costs. A firm whose revenue is equal to its economic costs is covering the opportunity cost of all of its factors of production, including capital. Economists would say that such a firm is earning normal profit, but not positive economic profit. It is earning a rate of return on capital just equal to the rate of return that an investor could expect to earn in an equivalently risky alternative investment (opportunity cost). Firms that are operating in a very competitive environment with no barriers to entry from other competitors can expect, in the long run, to be unable to earn a positive economic profit; the excess rate of return would attract entrants who would produce more output and ultimately drive the market price down to the level at which each firm is at best just earning a normal profit. This situation, of course, does not imply that the firm is earning zero accounting profit.

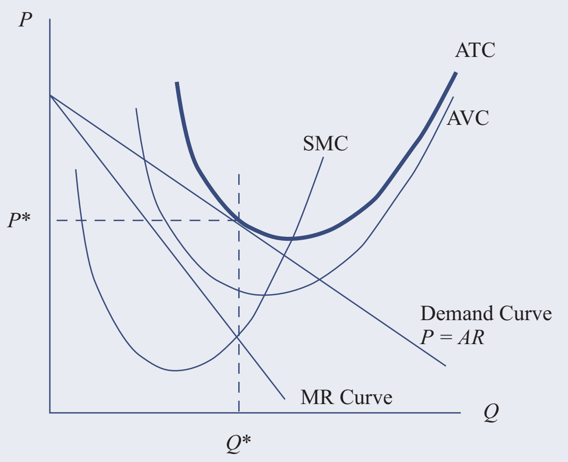
Exhibit 19 depicts the condition for both a firm under conditions of perfect competition (Panel A) and a monopolist (Panel B) in which the best each firm can do is to break even. Note that at the level of output at which SMC is equal to MR, price is just equal to ATC. Hence, economic profit is zero, and the firms are breaking even.

Exhibit 19 Examples of Firms under Perfect Competition and Monopolistic Firms That Can, at Best, Break Even

A. Perfect Competition



B. Monopolist



3.2.7 *The Shutdown Decision*

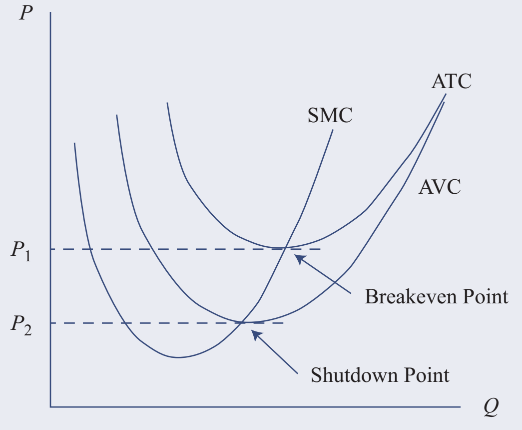
In the long run, if a firm cannot earn at least a zero economic profit, it will not operate because it is not covering the opportunity cost of all of its factors of production, labor, and capital. In the short run, however, a firm might find it advantageous to continue to operate even if it is not earning at least a zero economic profit. The discussion that follows addresses the decision to continue to operate and earn negative profit or shut down operations.

Recall that typically some or all of a firm's fixed costs are incurred regardless of whether the firm operates. The firm might have a lease on its building that it cannot avoid paying until the lease expires. In that case, the lease payment is a sunk cost: It cannot be avoided, no matter what the firm does. Sunk costs must be ignored in the decision to continue to operate in the short run. As long as the firm's revenues cover at least its variable cost, the firm is better off continuing to operate. If price is greater than AVC, the firm is not only covering all of its variable cost but also a portion of fixed cost.

For example, suppose a firm is producing 100 widgets and selling them at a price of €4 each. Obviously, its TR is €400 per time period. Suppose, also, that at that level of output, its ATC is €7, made up of AVC of €3.75 plus AFC of €3.25 per period. This firm is said to be earning negative economic profit (also referred to as **economic loss**, a condition in which revenues fall short of total opportunity cost) of €300 because its TC is €700 and its TR is only €400. Should this firm shut down immediately? If the fixed cost is unavoidable, then the firm is obligated to pay it whether it operates or not. The TFC is €325 (€3.25 per unit on 100 units). If it shuts down and earns zero revenue, then its variable cost would be zero but its losses would still equal the €325 of unavoidable fixed cost. If, however, it continued to operate, it could earn revenue of €400 that would cover its variable cost of €375 and contribute €25 toward the fixed costs. In other words, this firm would lose less by continuing to operate (€300) than by shutting down (€325).

In the long run, unless market price increases, this firm would exit the industry. But in the short run, it will continue to operate at a loss. Exhibit 20 depicts a firm under conditions of perfect competition facing three alternative market price ranges for its output. At any price above P_1, the firm can earn a positive profit and clearly should continue to operate. At a price below P_2, the minimum AVC, the firm could not even cover its variable cost and should shut down. At prices between P_2 and P_1, the firm should continue to operate in the short run because it is able to cover all of its variable cost and contribute something toward unavoidable fixed costs. Economists refer to the minimum AVC point as the **shutdown point** and the minimum ATC point as the **breakeven point**.

Exhibit 20 A Firm under Conditions of Perfect Competition Will Choose to Shut Down If Market Price Is Less Than Minimum AVC



EXAMPLE 5

Breakeven Analysis and Profit Maximization When the Firm Faces a Negatively Sloped Demand Curve under Imperfect Competition

Revenue and cost information for a future period is presented in Exhibit 21 for WR International, a newly formed corporation that engages in the manufacturing of low-cost, pre-fabricated dwelling units for urban housing markets in emerging economies. (Note that quantity increments are in blocks of 10 for a 250 change in price.) The firm has few competitors in a market setting of imperfect competition.

1 How many units must WR International sell to initially break even?
2 Where is the region of profitability?
3 At what point will the firm maximize profit? At what points are there economic losses?

Exhibit 21

Quantity (Q)	Price (P)	Total Revenue (TR)	Total Cost (TC)[a]	Profit
0	10,000	0	100,000	–100,000
10	9,750	97,500	170,000	–72,500
20	9,500	190,000	240,000	–50,000
30	9,250	277,500	300,000	–22,500
40	9,000	360,000	360,000	0
50	8,750	437,500	420,000	17,500
60	8,500	510,000	480,000	30,000
70	8,250	577,500	550,000	27,500
80	8,000	640,000	640,000	0
90	7,750	697,500	710,000	–12,500
100	7,500	750,000	800,000	–50,000

[a] Includes all opportunity costs

Solution to 1:

WR International will initially break even at 40 units of production, where TR and TC equal 360,000.

Solution to 2:

The region of profitability will range from greater than 40 units to less than 80 units. Any production quantity of less than 40 units and any quantity greater than 80 units will result in an economic loss.

Solution to 3:

Maximum profit of 30,000 will occur at 60 units. Lower profit will occur at any output level that is higher or lower than 60 units. From 0 units to less than 40 units and for quantities greater than 80 units, economic losses occur.

Given the relationships between TR, TVC, and TFC, Exhibit 22 summarizes the decisions to operate, shut down production, or exit the market in both the short run and the long run. The firm must cover its variable cost to remain in business in the short run; if TR cannot cover TVC, the firm shuts down production to minimize loss. The loss would be equal to the amount of fixed cost. If TVC exceeds TR in the long run, the firm will exit the market to avoid the loss associated with fixed cost at zero production. By exiting the market, the firm's investors do not suffer the erosion of their equity capital from economic losses. When TR is enough to cover TVC but not all of TFC, the firm can continue to produce in the short run but will be unable to maintain financial solvency in the long run.

Exhibit 22 Short Run and Long Run Decisions to Operate or Not

Revenue–Cost Relationship	Short-Run Decision	Long-Term Decision
TR = TC	Stay in market	Stay in market
TR = TVC but < TC	Stay in market	Exit market
TR < TVC	Shut down production	Exit market

EXAMPLE 6

Shutdown Analysis

For the most recent financial reporting period, a business domiciled in Ecuador (which recognizes the US dollar as an official currency) has revenue of $2 million and TC of $2.5 million, which are or can be broken down into TFC of $1 million and TVC of $1.5 million. The net loss on the firm's income statement is reported as $500,000 (ignoring tax implications). In prior periods, the firm had reported profits on its operations.

1 What decision should the firm make regarding operations over the short term?

2 What decision should the firm make regarding operations over the long term?

3 Assume the same business scenario except that revenue is now $1.3 million, which creates a net loss of $1.2 million. What decision should the firm make regarding operations in this case?

Solution to 1:

In the short run, the firm is able to cover all of its TVC but only half of its $1 million in TFC. If the business ceases to operate, its loss would be $1 million, the amount of TFC, whereas the net loss by operating would be minimized at $500,000. The firm should attempt to operate by negotiating special arrangements with creditors to buy time to return operations back to profitability.

Solution to 2:

If the revenue shortfall is expected to persist over time, the firm should cease operations, liquidate assets, and pay debts to the extent possible. Any residual for shareholders would decrease the longer the firm is allowed to operate unprofitably.

Solution to 3:

The firm would minimize loss at $1 million of TFC by shutting down. If the firm decided to continue to do business, the loss would increase to $1.2 million. Shareholders would save $200,000 in equity value by pursuing this option. Unquestionably, the business would have a rather short life expectancy if this loss situation were to continue.

When evaluating profitability, particularly of start-up firms and businesses using turnaround strategies, analysts should consider highlighting breakeven and shutdown points in their financial research. Identifying the unit sales levels at which the firm enters or leaves the production range for profitability and at which the firm can no longer function as a viable business entity provides invaluable insight for investment decisions.

3.3 Understanding Economies and Diseconomies of Scale

Rational behavior dictates that the firm select an operating size or scale that maximizes profit over any time frame. The time frame that defines the short run and long run for any firm is based on the ability of the firm to adjust the quantities of the fixed resources it uses. The short run is the time period during which at least one of the factors of production, such as technology, physical capital, and plant size, is fixed. The long run is defined as the time period during which all factors of production are variable. Additionally, in the long run, firms can enter or exit the market based on decisions regarding profitability. The long run is often referred to as the "planning horizon" in which the firm can choose the short-run position or optimal operating size that maximizes profit over time. The firm is always operating in the short run but planning in the long run.

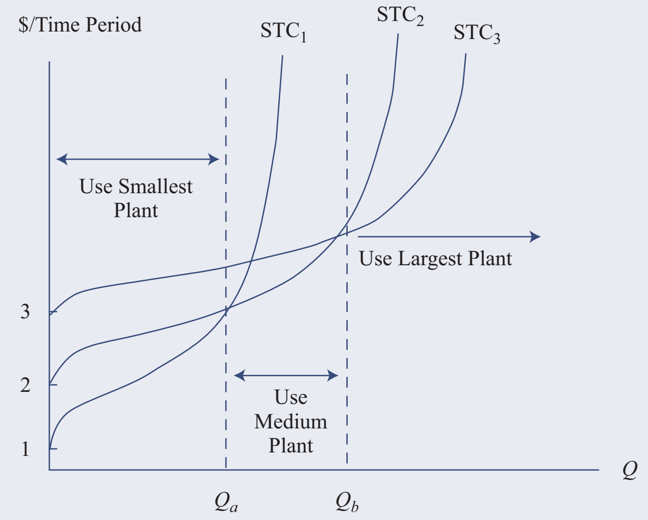
The time required for long-run adjustments varies by industry. For example, the long run for a small business using very little technology and physical capital may be less than a year, whereas for a capital-intensive firm, the long run may be more than a decade. Given enough time, however, all production factors are variable, which allows the firm to choose an operating size or plant capacity based on different technologies and physical capital. In this regard, costs and profits will differ between the short run and the long run.

3.3.1 *Short- and Long-Run Cost Curves*

Recall that when we addressed the short-run cost curves of the firm, we assumed that the capital input was held constant. That meant that the only way to vary output in the short run is to change the level of the variable input—in our case, labor. If the capital input—namely, plant and equipment—were to change, however, we would have an entirely new set of short-run cost curves, one for each level of capital input.

The short-run total cost includes all the inputs—labor and capital—the firm is using to produce output. For reasons discussed earlier, the typical short-run total cost (STC) curve might rise with output, first at a decreasing rate because of specialization economies and then at an increasing rate, reflecting the law of diminishing marginal returns to labor. Total fixed cost (the quantity of capital input multiplied by the rental rate on capital) determines the vertical intercept of the STC curve. At higher levels of fixed input, TFC is greater but the production capacity of the firm is also greater. Exhibit 23 shows three different STC curves for the same technology but using three distinct levels of capital input—Points 1, 2, and 3 on the vertical axis.

Exhibit 23 Short-Run Total Cost Curves for Various Plant Sizes



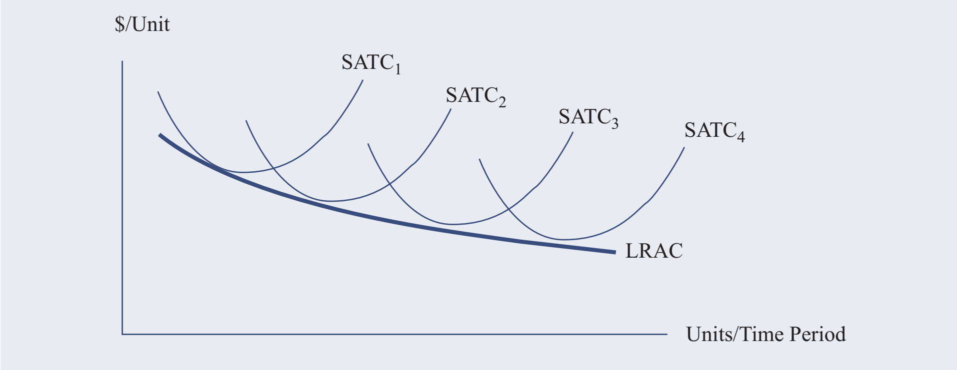
Plant Size 1 is the smallest and, of course, has the lowest fixed cost; hence, its STC_1 curve has the lowest vertical intercept. But note that STC_1 begins to rise more steeply with output, reflecting the lower plant capacity. Plant Size 3 is the largest of the three and reflects that size with both a higher fixed cost and a lower slope at any level of output. If a firm decided to produce an output between zero and Q_a, it would plan on building Plant Size 1 because for any output level in that range, its cost is less than it would be for Plant Size 2 or 3. Accordingly, if the firm were planning to produce output greater than Q_b, it would choose Plant Size 3 because its cost for any of those levels of output would be lower than for Plant Size 1 or 2. And of course, Plant Size 2 would be chosen for output levels between Q_a and Q_b. The long-run total cost curve is derived from the lowest level of STC for each level of output because in the long run, the firm is free to choose which plant size it will operate. This curve is called an "envelope curve." In essence, this curve envelopes—encompasses—all possible combinations of technology, plant size, and physical capital.

For each STC curve, there is also a corresponding **short-run average total cost** (SATC) curve and a corresponding **long-run average total cost** (LRAC) curve, the envelope curve of all possible short-run average total cost curves. The shape of the LRAC curve reflects an important concept called **economies of scale** and **diseconomies of scale**.

3.3.2 *Defining Economies of Scale and Diseconomies of Scale*

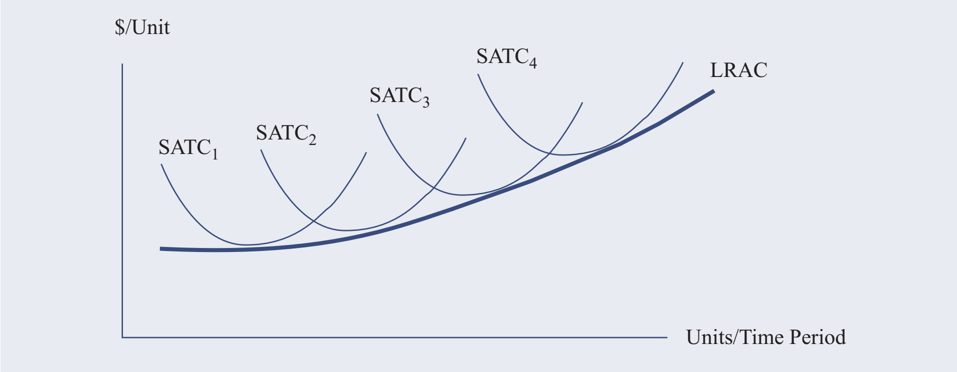
When a firm increases all of its inputs in order to increase its level of output (obviously, a long-run concept), it is said to *scale up* its production. *Scaling down* is the reverse—decreasing all of its inputs in order to produce less in the long run. Economies of scale occur if, as the firm increases its output, cost per unit of production falls. Graphically, this definition translates into a LRAC curve with a negative slope. Exhibit 24 depicts several SATC curves, one for each plant size, and the LRAC curve representing economies of scale.

Exhibit 24 Short-run Average Total Cost Curves for Various Plant Sizes and Their Envelope Curve, LRAC: Economies of Scale



Diseconomies of scale occur if cost per unit rises as output increases. Graphically, diseconomies of scale translate into an LRAC curve with a positive slope. Exhibit 25 depicts several SATC curves, one for each plant size, and their envelope curve, the LRAC curve, representing diseconomies of scale.

Exhibit 25 Short-run Average Total Cost Curves for Various Plant Sizes and Their Envelope Curve, LRAC: Diseconomies of Scale



As the firm grows in size, economies of scale and a lower ATC can result from the following factors:

- **Increasing returns to scale**, which is when a production process allows for increases in output that are proportionately larger than the increase in inputs.
- Having a division of labor and management in a large firm with numerous workers, which allows each worker to specialize in one task rather than perform many duties, as in the case of a small business (as such, workers in a large firm become more proficient at their jobs).
- Being able to afford more expensive, yet more efficient equipment and to adapt the latest in technology that increases productivity.
- Effectively reducing waste and lowering costs through marketable byproducts, less energy consumption, and enhanced quality control.
- Making better use of market information and knowledge for more effective managerial decision making.
- Obtaining discounted prices on resources when buying in larger quantities.

A classic example of a business that realizes economies of scale through greater physical capital investment is an electric utility. By expanding output capacity to accommodate a larger customer base, the utility company's per-unit cost will decline. Economies of scale help explain why electric utilities have naturally evolved from localized entities to regional and multi-region enterprises. Wal-Mart is an example of a business that uses bulk purchasing power to obtain deep discounts from suppliers to keep costs and prices low. Wal-Mart also uses the latest technology to monitor point-of-sale transactions to gather timely market information to respond to changes in customer buying behavior, which leads to economies of scale through lower distribution and inventory costs.

The factors that can lead to diseconomies of scale, inefficiencies, and rising costs when a firm increases in size include the following:

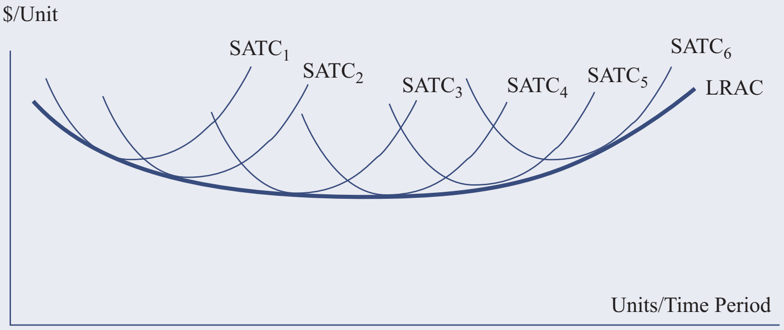
- **Decreasing returns to scale**, which is when a production process leads to increases in output that are proportionately smaller than the increase in inputs.
- Being so large that it cannot be properly managed.
- Overlapping and duplication of business functions and product lines.
- Higher resource prices because of supply constraints when buying inputs in large quantities.

Before its restructuring, General Motors (GM) was an example of a business that had realized diseconomies of scale by becoming too large. Scale diseconomies occurred through product overlap and duplication (i.e., similar or identical automobile models), and the fixed cost for these models was not spread over a large volume of output. (Recently, the company has decided to discontinue various low-volume product models that overlapped with other models.) GM had numerous manufacturing plants throughout the world and sold vehicles in more than a hundred countries. Given this geographical dispersion in production and sales, the company had communication and management coordination problems, which resulted in higher costs. Also, GM had significantly higher labor costs than its competitors. As the largest producer in the market, it had been a target of labor unions for higher compensation and benefits packages relative to other firms.

Economies and diseconomies of scale can occur at the same time; the impact on long-run average total cost (LRAC) depends on which dominates. If economies of scale dominate, LRAC decreases with increases in output. The reverse holds true when diseconomies of scale prevail. There may be a range of output over which LRAC falls (economies of scale) and then a range over which LRAC might be constant, followed by a range over which diseconomies of scale prevail, as depicted in Exhibit 26.

The minimum point on the LRAC curve is referred to as the **minimum efficient scale**. The minimum efficient scale is the optimal firm size under perfect competition over the long run. Theoretically, perfect competition forces the firm to operate at the minimum point on the LRAC curve because the market price will be established at this level over the long run. If the firm is not operating at this least-cost point, its long-term viability will be threatened.

Exhibit 26 LRAC Can Exhibit Economies and Diseconomies of Scale



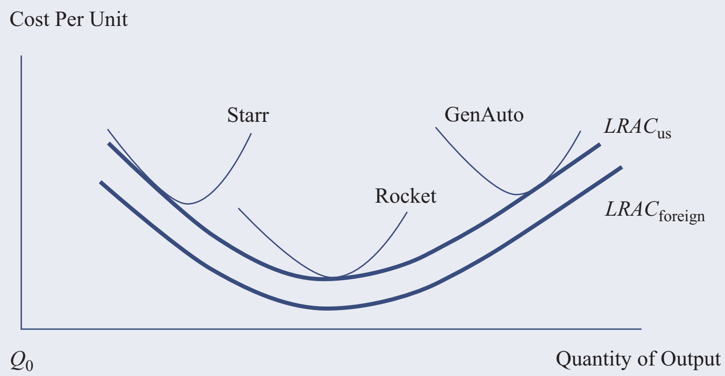
EXAMPLE 7

Long-Run Average Total Cost Curve

Exhibit 27 displays the long-run average total cost curve ($LRAC_{US}$) and the short-run average total cost curves for three hypothetical US-based automobile manufacturers—Starr Vehicles (Starr), Rocket Sports Cars (Rocket), and General Auto (GenAuto). The LRAC curve for foreign-owned automobile companies that compete in the US auto market ($LRAC_{foreign}$) is also indicated in the graph. (The market structure implicit in the exhibit is imperfect competition.)

To what extent are the cost relationships depicted in Exhibit 27 useful for an economic and financial analysis of the three US-based auto firms?

Exhibit 27



Solution:

First, it is observable that the foreign auto companies have a lower LRAC compared with that of the US automobile manufacturers. This competitive position places the US firms at a cost—and possibly, pricing—disadvantage in the market, with the potential to lose market share to the lower-cost foreign competitors. Second, only Rocket operates at the minimum point of the $LRAC_{US}$, whereas GenAuto is situated in the region of diseconomies of scale and Starr is positioned in the economies of scale portion of the curve. To become more efficient and competitive, GenAuto needs to downsize and restructure, which means moving down the $LRAC_{US}$ curve to a smaller, yet lower-cost production volume. In Contrast, Starr has to grow in size to become more efficient and competitive by lowering per-unit costs.

From a long-term investment prospective and given its cost advantage, Rocket has the potential to create more investment value relative to GenAuto and Starr. Over the long run, if GenAuto and Starr can lower their ATC, they will become more attractive to investors. But if any of the three US auto companies cannot match the cost competitiveness of the foreign firms, they may be driven from the market. In the long run, the lower-cost foreign automakers pose a severe competitive challenge to the survival of the US manufacturers and their ability to maintain and grow shareholders' wealth.

SUMMARY

This reading addressed several important concepts that extend the basic market model of demand and supply to assist the analyst in assessing a firm's breakeven and shutdown points of production. Demand concepts covered include own-price elasticity of demand, cross-price elasticity of demand, and income elasticity of demand. Supply concepts covered include total, average, and marginal product of labor; total, variable, and marginal cost of labor; and total and marginal revenue. These concepts are used to calculate the breakeven and shutdown points of production.

- Elasticity of demand is a measure of how sensitive quantity demanded is to changes in various variables.
- Own-price elasticity of demand is the ratio of percentage change in quantity demanded to percentage change in a good or service's own price.
- If own-price elasticity of demand is greater than one in absolute terms, demand is elastic and a decline in price will result in higher total expenditure on that good.
- If own-price elasticity of demand is less than one in absolute terms, demand is inelastic and a decline in price will result in a lower total expenditure on that good.
- If own-price elasticity of demand is equal to negative one, demand is unit, or unitary, elastic and total expenditure on that good is independent of price.
- Own-price elasticity of demand will almost always be negative.
- Income elasticity of demand is the ratio of the percentage change in quantity demanded to the percentage change in consumer income.
- Demand is negatively sloped because of either the substitution effect or the income effect.
- The substitution effect is the phenomenon in which, as a good's price falls, more of this good is substituted for other, more expensive goods.
- The income effect is the phenomenon in which, as a good's price falls, real income rises and, if this good is normal, more of it will be purchased.
- If the good is inferior, the income effect will partially or fully offset the substitution effect.
- There are two exceptions to the law of demand: Giffen goods and Veblen goods.

- Giffen goods are highly inferior and make up a large portion of the consumer budget. As price falls, the substitution effect tends to cause more of the good to be consumed, but the highly negative income effect overwhelms the substitution effect. Demand curves for Giffen goods are positively sloped.
- Veblen goods are highly valued high-priced “status” goods; consumers may tend to buy more of a good if its price rises.
- If income elasticity of demand is positive, the good is a normal good. If income elasticity of demand is negative, the good is an inferior good.
- Cross-price elasticity of demand is the ratio of the percentage change in quantity demanded of one good to the percentage change in the price of a related good.
- If cross-price elasticity between two goods is positive, they are substitutes, and if cross-price elasticity between two goods is negative, they are complements.
- The law of demand states that a decrease in price will cause an increase in quantity demanded.
- Total product of labor is a short-run concept that is the total quantity that is able to be produced for each level of labor input, holding all other inputs constant.
- Average product of labor (APL) is the total product of labor divided by number of labor hours.
- Marginal product of labor (MP_L) is the change in total product divided by the change in labor hours. MP_L might rise as more labor is added to a fixed amount of capital.
- The law of diminishing returns dictates that additional output must fall as more and more labor is added to a fixed amount of capital.
- Production costs increase as input prices rise and fall as inputs become more productive.
- Short-run total cost (STC) is the total expenditure on fixed capital plus the total expenditure on labor.
- Short-run marginal cost (SMC) equals the ratio of wage to marginal product of labor (MP_L).
- Average variable cost (AVC) is the ratio of wage to average product of labor (APL).
- Average total cost (ATC) is total cost (TC) divided by the number of units produced.
- Revenue is price times quantity sold.
- Marginal revenue (MR) is the ratio of change in revenue to change in output.
- Firms under conditions of perfect competition have no pricing power and, therefore, face a perfectly horizontal demand curve at the market price. For firms under conditions of perfect competition, price is identical to marginal revenue (MR).
- Firms under conditions of imperfect competition face a negatively sloped demand curve and have pricing power. For firms under conditions of imperfect competition, marginal revenue (MR) is less than price.
- Economic profit equals total revenue (TR) minus total economic cost, whereas accounting profit equals TR minus total accounting cost.
- Economic cost takes into account the total opportunity cost of all factors of production.
- Opportunity cost is the next best alternative forgone in making a decision.

- Maximum economic profit requires that (1) marginal revenue (MR) equals marginal cost (MC) and (2) MC not be falling with output.
- The breakeven point occurs when total revenue (TR) equals total cost (TC), otherwise stated as the output quantity at which average total cost (ATC) equals price.
- Shutdown occurs when a firm is better off not operating than continuing to operate.
- If all fixed costs are sunk costs, then shutdown occurs when the market price falls below minimum average variable cost. After shutdown, the firm incurs only fixed costs and loses less money than it would operating at a price that does not cover variable costs.
- In the short run, it may be rational for a firm to continue to operate while earning negative economic profit if some unavoidable fixed costs are covered.
- Economies of scale is defined as decreasing long-run cost per unit as output increases. Diseconomies of scale is defined as increasing long-run cost per unit as output increases.
- Long-run average total cost is the cost of production per unit of output under conditions in which all inputs are variable.
- Specialization efficiencies and bargaining power in input price can lead to economies of scale.
- Bureaucratic and communication breakdowns and bottlenecks that raise input prices can lead to diseconomies of scale.
- The minimum point on the long-run average total cost curve defines the minimum efficient scale for the firm.

PRACTICE PROBLEMS

1 If the price elasticity coefficient of the demand curve for paper clips is equal to –1, demand is:

A elastic.

B inelastic.

C unit elastic.

2 The demand for membership at a local health club is determined by the following equation:

$$Q_{hm}^{d} = 400 - 5P_{hm}$$

where Q_{hm}^{d} is the number of health club members and P_{hm} is the price of membership. If the price of health club membership is $35, the price elasticity of demand is *closest* to:

A –0.778.

B –0.500.

C –0.438.

3 Price elasticity of demand for a good will *most likely* be greater if:

A there are no substitutes for the good.

B consumers consider the good as discretionary.

C consumers spend a small portion of their budget on the good.

4 If the income elasticity of demand for a product is –0.6, a:

A 1% increase in income will result in a 0.6% increase in demand.

B 1% increase in income will result in a 0.6% decrease in demand.

C 0.6% increase in income will result in a 1% decrease in demand.

5 An individual's demand for onions is given by the following equation:

$$Q_{o}^{d} = 3 - 0.05P_{o} + 0.009I - 0.16P_{t}$$

where Q_{o}^{d} is the number of onions demanded, P_o is the price per pound of onions, I is the household income, and P_t is the price per pound of tomatoes. If the price of onions is $1.25, household income is $2,500, and the price of tomatoes is $3.75, the cross-price elasticity of demand for onions with respect to the price of tomatoes is *closest* to:

A –1.0597.

B –0.0242.

C –0.0081.

6 Movement along the demand curve for good *X* occurs due to a change in:

A income.

B the price of good *X*.

C the price of a substitute for good *X*.

7 A wireless phone manufacturer introduced a next-generation phone that received a high level of positive publicity. Despite running several high-speed production assembly lines, the manufacturer is still falling short in meeting demand for the phone nine months after introduction. Which of the following statements is the *most* plausible explanation for the demand/supply imbalance?

- **A** The phone price is low relative to the equilibrium price.
- **B** Competitors introduced next-generation phones at a similar price.
- **C** Consumer incomes grew faster than the manufacturer anticipated.

The following information relates to Questions 8–11

The market demand function for four-year private universities is given by the equation

$$Q_{pr}^{d} = 84 - 3.1P_{pr} + 0.8I + 0.9P_{pu}$$

where Q_{pr}^{d} is the number of applicants to private universities per year in thousands, P_{pr} is the average price of private universities (in thousands of USD), I is the household monthly income (in thousands of USD), and P_{pu} is the average price of public (government-supported) universities (in thousands of USD). Assume that P_{pr} is equal to 38, I is equal to 100, and P_{pu} is equal to 18.

8 The price elasticity of demand for private universities is *closest* to:

- **A** –3.1.
- **B** –1.9.
- **C** 0.6.

9 The income elasticity of demand for private universities is *closest* to:

- **A** 0.5.
- **B** 0.8.
- **C** 1.3.

10 The cross-price elasticity of demand for private universities with respect to the price of public universities is *closest* to:

- **A** 0.3.
- **B** 3.1.
- **C** 3.9.

11 If the cross-price elasticity between two goods is negative, the two goods are classified as:

- **A** normal.
- **B** substitutes.
- **C** complements.

---

12 In the case of a normal good with a decrease in own price, which of the following statements is *most likely* true?

- **A** Both the substitution and income effects lead to an increase in the quantity purchased.

B The substitution effect leads to an increase in the quantity purchased, while the income effect has no impact.

C The substitution effect leads to an increase in the quantity purchased, while the income effect leads to a decrease.

13 For a Giffen good, the:

A demand curve is positively sloped.

B substitution effect overwhelms the income effect.

C income and substitution effects are in the same direction.

14 Normal profit is best described as:

A zero economic profit.

B total revenue minus all explicit costs.

C the sum of accounting profit plus economic profit.

15 A company plans to hire additional factory employees. In the short run, marginal returns are most likely to decrease if:

A the factory is operating at full capacity.

B the factory is experiencing a labor shortage.

C workers are required to multitask and share duties.

16 The production relationship between the number of machine hours and total product for a company is presented below.

Machine Hours	Total Product	Average Product
1	3	3.00
2	8	4.00
3	14	4.67
4	19	4.75
5	21	4.20

Diminishing marginal returns first occur beyond machine hour:

A 3.

B 4.

C 5.

17 The marketing director for a Swiss specialty equipment manufacturer estimates the firm can sell 200 units and earn total revenue of CHF500,000. However, if 250 units are sold, revenue will total CHF600,000. The marginal revenue per unit associated with marketing 250 units instead of 200 units is *closest* to:

A CHF 2,000.

B CHF 2,400.

C CHF 2,500.

18 An agricultural firm operating in a perfectly competitive market supplies wheat to manufacturers of consumer food products and animal feeds. If the firm were able to expand its production and unit sales by 10% the *most likely* result would be:

A a 10% increase in total revenue.

B a 10% increase in average revenue.

C an increase in total revenue of less than 10%.

19 An operator of a ski resort is considering offering price reductions on weekday ski passes. At the normal price of €50 per day, 300 customers are expected to buy passes each weekday. At a discounted price of €40 per day 450 customers are expected to buy passes each weekday. The marginal revenue per customer earned from offering the discounted price is *closest* to:

A €20.

B €40.

C €50.

20 The marginal revenue per unit sold for a firm doing business under conditions of perfect competition will *most likely* be:

A equal to average revenue.

B less than average revenue.

C greater than average revenue.

The following information relates to Questions 21–23

A firm's director of operations gathers the following information about the firm's cost structure at different levels of output:

Exhibit 1

Quantity (*Q*)	Total Fixed Cost (*TFC*)	Total Variable Cost (*TVC*)
0	200	0
1	200	100
2	200	150
3	200	200
4	200	240
5	200	320

21 Refer to the data in Exhibit 1. When quantity produced is equal to 4 units, the average fixed cost (*AFC*) is *closest* to:

A 50.

B 60.

C 110.

22 Refer to the data in Exhibit 1. When the firm increases production from 4 to 5 units, the marginal cost (*MC*) is *closest* to:

A 40.

B 64.

C 80.

23 Refer to the data in Exhibit 1. The level of unit production resulting in the lowest average total cost (*ATC*) is *closest* to:

A 3.

B 4.
C 5.

---

24 The short-term breakeven point of production for a firm operating under perfect competition will *most likely* occur when:
A price is equal to average total cost.
B marginal revenue is equal to marginal cost.
C marginal revenue is equal to average variable costs.

25 The short-term shutdown point of production for a firm operating under perfect competition will *most likely* occur when:
A price is equal to average total cost.
B marginal revenue is equal to marginal cost.
C marginal revenue is equal to average variable costs.

26 Under conditions of perfect competition, a company will break even when market price is equal to the minimum point of the:
A average total cost curve.
B average variable cost curve.
C short-run marginal cost curve.

27 A company will shut down production in the short run if total revenue is less than total:
A fixed costs.
B variable costs.
C opportunity costs.

28 A company has total variable costs of $4 million and fixed costs of $3 million. Based on this information, the company will stay in the market in the long term if total revenue is at least:
A $3.0 million.
B $4.5 million.
C $7.0 million.

29 When total revenue is greater than total variable costs but less than total costs, in the short term a firm will *most likely*:
A exit the market.
B stay in the market.
C shut down production.

30 A profit maximum is *least likely* to occur when:
A average total cost is minimized.
B marginal revenue equals marginal cost.
C the difference between total revenue and total cost is maximized.

31 A firm that increases its quantity produced without any change in per-unit cost is experiencing:
A economies of scale.
B diseconomies of scale.
C constant returns to scale.

32 A company is experiencing economies of scale when:

A cost per unit increases as output increases.

B it is operating at a point on the LRAC curve where the slope is negative.

C It is operating beyond the minimum point on the long-run average total cost curve.

33 Diseconomies of scale *most likely* result from:

A specialization in the labor force.

B overlap of business functions and product lines.

C discounted prices on resources when buying in larger quantities.

34 A firm is operating beyond minimum efficient scale in a perfectly competitive industry. To maintain long-term viability the *most likely* course of action for the firm is to:

A operate at the current level of production.

B increase its level of production to gain economies of scale.

C decrease its level of production to the minimum point on the long-run average total cost curve.

35 Under conditions of perfect competition, in the long run firms will *most likely* earn:

A normal profits.

B positive economic profits.

C negative economic profits.

The following information relates to Questions 36 and 37

The manager of a small manufacturing firm gathers the following information about the firm's labor utilization and production:

Exhibit 2

Labor (*L*)	Total Product (*TP*)
0	0
1	150
2	320
3	510
4	660
5	800

36 Refer to the data in Exhibit 2. The number of workers resulting in the highest level of average product of labor is *closest* to:

A 3.

B 4.

C 5.

37 Refer to the data in Exhibit 2. The marginal product of labor demonstrates increasing returns for the firm if the number of workers is *closest* to but not more than:

A 2.

B 3.

C 4.

SOLUTIONS

1 C is correct. When the price elasticity of demand coefficient is –1, demand is said to be unit elastic, or unitary elastic.

2 A is correct. Inserting the price of $35 into the demand function, quantity demanded is calculated as

$$Q_{hm}^{d} = 400 - 5(35) = 225$$

At a price of $35 per health club membership, the elasticity of demand is

$$\text{Price elasticity of demand} = \left(\Delta Q_{hm}^{d} / \Delta P_{hm}\right) \times \left(P_{hm} / Q_{hm}^{d}\right)$$

$$\text{Price elasticity of demand} = -5 \times (35/225) = -0.778$$

3 B is correct. Price elasticity of demand is likely to be greater for items that are seen as optional or discretionary.

4 B is correct. Income elasticity is a measure of how sensitive quantity demanded is to a change in income. If the income elasticity of demand for the product is –0.6, whenever income increases by 1%, the quantity demanded of the product at each price decreases by 0.6%. Consequently, as income rises, consumers will purchase less of the product.

5 B is correct. The cross-price elasticity of demand measures the responsiveness of the demand for onions in response to a change in the price of tomatoes. From the demand function equation:

$$Q_{o}^{d} = 3 - 0.05P_{o} + 0.009I - 0.16P_{t}$$

$$Q_{o}^{d} = 3 - 0.05(1.25) + 0.009(2{,}500) - 0.16(3.75) = 24.8375$$

At a price of onions of $1.25 and a price of tomatoes of $3.75, the cross-price elasticity of demand is calculated as follows:

$$\text{Cross-price elasticity of demand} = \left(\Delta Q_{o}^{d} / \Delta P_{t}\right) \times \left(P_{t} / Q_{o}^{d}\right)$$

$$\text{Cross-price elasticity of demand} = -0.16 \times (3.75/24.8375) = -0.0242$$

6 B is correct. The demand curve shows quantity demanded as a function of own price only.

7 A is correct. The situation described is one of excess demand because, in order for markets to clear at the given level of quantity supplied, the company would need to raise prices.

8 B is correct. From the demand function:

Solve for Q_{pr}^{d}:

$$\Delta Q_{pr}^{d} / \Delta P_{pr} = -3.1 \text{ (the coefficient in front of own price)}$$

$$\begin{aligned} Q_{pr}^{d} &= 84 - 3.1P_{pr} + 0.8I + 0.9P_{pu} \\ &= 84 - 3.1(38) + 0.8(100) + 0.9(18) \\ &= 62.4 \end{aligned}$$

At P_{pr} = 38,

$$\text{price elasticity of demand} = \left(\Delta Q^d_{pr} / \Delta P_{pr}\right)\left(P_{pr} / Q^d_{pr}\right)$$
$$= (-3.1)(38/62.4)$$
$$= -1.9$$

9 C is correct. From the demand function:

Solve for Q^d_{pr}:

$$\Delta Q^d_{pr} / \Delta I = 0.8 \text{ (coefficient in front of the income variable)}$$

$$Q^d_{pr} = 84 - 3.1P_{pr} + 0.8I + 0.9P_{pu}$$
$$= 84 - 3.1(38) + 0.8(100) + 0.9(18)$$
$$= 62.4$$

At I = 100,

$$\text{the income elasticity of demand} = \left(\Delta Q^d_{pr} / \Delta I\right)\left(I / Q^d_{pr}\right)$$
$$= (0.8)(100/62.4)$$
$$= 1.3$$

10 A is correct. From the demand function:

Solve for Q^d_{pr}:

$$\Delta Q^d_{pr} / \Delta P_{pu} = 0.9 \text{ (the coefficient in front of } P_{pu})$$

$$Q^d_{pr} = 84 - 3.1P_{pr} + 0.8I + 0.9P_{pu}$$
$$= 84 - 3.1(38) + 0.8(100) + 0.9(18)$$
$$= 62.4$$

At P = 38, and P_{pu} = 18,

$$\text{the cross-price elasticity of demand} = \left(\Delta Q^d_{pr} / \Delta P_{pu}\right)\left(P_{pu} / Q^d_{pr}\right)$$
$$= (0.9)(18/62.4)$$
$$= 0.3$$

11 C is correct. With complements, consumption goes up or down together. With a negative cross-price elasticity, as the price of one good goes up, the demand for both falls.

12 A is correct. In the case of normal goods, the income and substitution effects are reinforcing, leading to an increase in the amount purchased after a drop in price.

13 A is correct. The income effect overwhelms the substitution effect such that an increase in the price of the good results in greater demand for the good, resulting in a positively sloped demand curve.

14 A is correct. Normal profit is the level of accounting profit such that implicit opportunity costs are just covered; thus, it is equal to a level of accounting profit such that economic profit is zero.

15 A is correct. The law of diminishing returns occurs in the short run when additional output falls as more and more labor is added to a fixed amount of capital. When a factory is operating at full capacity, adding additional employees will not increase production because the physical plant is already 100% employed. More labor hours will add to costs without adding to output, thus resulting in diminishing marginal returns.

16 A is correct. Diminishing marginal returns occur when the marginal product of a resource decreases as additional units of that input are employed. Marginal product, which is the additional output resulting from using one more unit of input, is presented below.

Machine Hours	Total Product	Average Product	Marginal Product
1	3	3.00	3
2	8	4.00	5
3	14	4.67	6
4	19	4.75	5
5	21	4.20	2

The marginal product of the third machine hour is 6 and declines thereafter. Consequently, diminishing marginal returns are first evident beyond three machine hours.

17 A is correct. Marginal revenue per unit is defined as the change in total revenue divided by the change in quantity sold. MR = ΔTR ÷ ΔQ. In this case, change in total revenue equals CHF100,000, and change in total units sold equals 50. CHF100,000 ÷ 50 = CHF2,000.

18 A is correct. In a perfectly competitive market, an increase in supply by a single firm will not affect price. Therefore, an increase in units sold by the firm will be matched proportionately by an increase in revenue.

19 A is correct. Marginal revenue per unit is defined as the change in total revenues divided by the change in quantity sold. MR = ΔTR ÷ ΔQ. In this case, change in total revenue per day equals €3,000 [(450 × €40) – (300 × €50)], and change in units sold equals 150 (450 – 300). €3,000 ÷ 150 = €20.

20 A is correct. Under perfect competition, a firm is a price taker at any quantity supplied to the market, and AR = MR = Price.

21 A is correct. Average fixed cost is equal to total fixed cost divided by quantity produced: AFC = TFC/Q = 200/4 = 50.

22 C is correct. Marginal cost is equal to the change in total cost divided by the change in quantity produced. MC = ΔTC/ΔQ = 80/1 = 80.

23 C is correct. Average total cost is equal to total cost divided by quantity produced. At 5 units produced the average total cost is 104. ATC = TC/Q = 520/5 = 104.

24 A is correct. Under perfect competition, price equals marginal revenue. A firm breaks even when marginal revenue equals average total cost.

25 C is correct. The firm should shut down production when marginal revenue is less than average variable cost.

26 A is correct. A company is said to break even if its total revenue is equal to its total cost. Under conditions of perfect competition, a company will break even when market price is equal to the minimum point of the average total cost curve.

27 B is correct. A company will shut down production in the short run when total revenue is below total variable costs.

28 C is correct. A company will stay in the market in the long term if total revenue is equal to or greater than total cost. Because total costs are \$7 million (\$4 million variable costs and \$3 million fixed costs), the company will stay in the market in the long term if total revenue equals at least \$7 million.

29 B is correct. When total revenue is enough to cover variable costs but not total fixed costs in full, the firm can survive in the short run but would be unable to maintain financial solvency in the long run.

30 A is correct. The quantity at which average total cost is minimized does not necessarily correspond to a profit maximum.

31 C is correct. Output increases in the same proportion as input increases occur at constant returns to scale.

32 B is correct. Economies of scale occur if, as the firm increases output, cost per unit of production falls. Graphically, this definition translates into a long-run average cost curve (LRAC) with a negative slope.

33 B is correct. As the firm increases output, diseconomies of scale and higher average total costs can result when there is overlap and duplication of business functions and product lines.

34 C is correct. The firm operating at greater than long-run efficient scale is subject to diseconomies of scale. It should plan to decrease its level of production.

35 A is correct. Competition should drive prices down to long-run marginal cost, resulting in only normal profits being earned.

36 A is correct. Three workers produce the highest average product equal to 170. AP = 510/3 = 170.

37 B is correct. Marginal product is equal to the change in total product divided by the change in labor. The increase in MP from 2 to 3 workers is 190: MP = $\Delta TP/\Delta L = (510 - 320)/(3 - 2) = 190/1 = 190$.