
2. *Supply and Demand: Consumer Demand (Prerequisite)*

Candidate Note: Candidates are expected to be familiar with the material covered in this prerequisite economics reading.

Learning Objectives

This summary includes a review and an analysis of the principles set forth by CFA Institute. Upon review of this summary, you should be able to:

- ❖ *Explain consumer choice and utility theoriespg. 32*
- ❖ *Understand how indifference curves, opportunity sets, and budget constraints apply to decision makingpg. 32*
- ❖ *Explain budget constraints, including their calculationpg. 34*
- ❖ *Find the equilibrium bundle of goods for a consumer, using utility analysispg. 36*

Consumer Choice and Utility Theories

Learning Objective: Explain consumer choice and utility theories.

How Consumer Choice Affects the Law of Demand

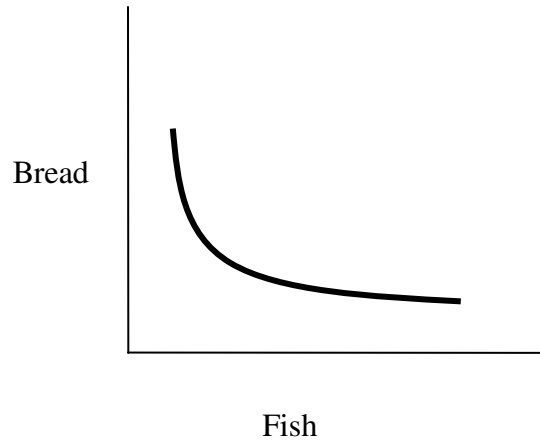
- Scarcity necessitates rationing; a limited amount of desired resources necessitates a mechanism for distributing such resources. Prices serve to “ration” or allocate scarce goods and resources to those people who are willing to pay the highest price.
- Scarcity results in competition: Scarce resources promote competition among individuals since all demand cannot be satisfied. When the rationing criterion is price, able individuals engage in income-generating activities that enhance their ability to pay the price. A different rationing mechanism (i.e., politics) encourages other types of behavior.

A consumer’s **budget constraint** is *the ability to buy goods and services, given income*. A consumer would like to consume a particular consumption basket, a bundle of goods and services. Utility theory assumes consumers know what they want, in other words, they can determine which bundle of goods and services is preferred. The theory assumes preferences are transitive, in other words, if you prefer bundle A to bundle B, and you prefer bundle B to bundle C, then you must also prefer bundle A to bundle C. The theory also assumes that for at least one good in the bundle, the consumer will always want more of one good. When “more is worse,” then the “good” is removal of the “bad” item. For example, more pollution is bad, and consumers pay for the removal of pollution. In some cases, some is good but too much is bad. Risk is an example of this - consumers may be willing to take some risks, but at some point have too much risk and will pay for increased certainty.

Indifference Curves

Learning Objective: Understand how indifference curves, opportunity sets, and budget constraints apply to decision making.

Indifference curve - *A convex curve that differentiates between consumption bundles that are more or less preferred by an individual. The points on the curve represent combinations of goods that are equally preferred by the individual.*



Indifference Curve: Bread and Fish

Characteristics of Indifference Curves

- Consumers prefer more rather than less. Bundles on indifference curves lying further from the origin are always preferred.
- Goods may be substituted for each other. Indifference curves have a negative and decreasing slope (i.e., as the consumption of one good falls, in order to maintain the same level of utility, consumption of the other good must rise).
- As a good is consumed more intensively, its valuation declines. Therefore, indifference curves are convex when viewed from below. The slope of the indifference curve represents the willingness of the individual to substitute one good for the other. The slope is thus called the marginal rate of substitution and is defined below.
- Individuals can compare any two bundles of goods anywhere on the diagram. One bundle will either be preferred to the other or they will be equally liked.
- Indifference curves can never cross. If they did, rational ordering would be violated.

Marginal rate of substitution (MRS): *Changing the consumption level of one good is just enough to offset the change in the consumption of another good without causing the individual to move to another indifference curve.* MRS is equal to the slope of the indifference curve.

Consumption opportunity constraint: The constraint that *divides attainable consumption bundles from unattainable consumption bundles.* In a barter economy, consumer satisfaction will be maximized when an individual's indifference curve is tangent to the market constraint curve (the rate at which goods can be exchanged). *In a money income economy, the consumption opportunity constraint* is usually called the **budget constraint**.

Real income is *the set of consumption possibilities the consumer has - what goods and services can be purchased with the consumer's fixed income.* If a consumer has a fixed income, reducing the price of a good means real income rises. If a consumer's monthly car payment declines, the consumer can spend more on other goods.

Producers perform the following functions:

- organize productive inputs
- transform and combine inputs into goods
- sell the final products to consumers for a price

The cost of production to the producer equals the product's opportunity cost. The opportunity cost represents the value of those things given up by society to produce the product.

Profit is *excess sales revenue relative to the cost of production, which includes the opportunity cost of all resources*. Profit accrues only when the market value of the good produced is greater than the sum of the costs of the individual resources used to produce it. **Losses** result if the sales revenue derived from a project is insufficient to pay the opportunity cost of the resources used.

Successful producers convert and rearrange resources in a way that increases their value. That is, the final product is worth more than the combined opportunity costs of its inputs.

Budget Constraints

Learning Objective: Explain budget constraints, including their calculation.

Suppose a consumer has a given amount of income, I . For simplicity, assume a one-period model in which he can only purchase two goods, A and B, at prices P_A and P_B . His income or budget constraint is:

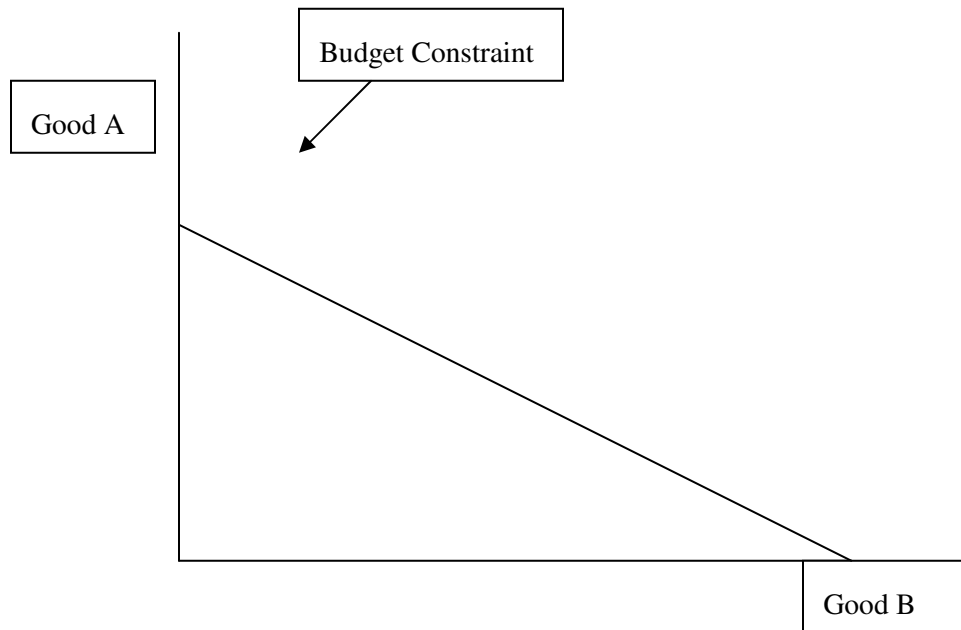
$$P_A \times \text{Quantity}_A + P_B \times \text{Quantity}_B = I$$

Rearranging terms:

$$\text{Quantity}_A = \frac{I}{P_A} - \frac{P_B}{P_A} \times \text{Quantity}_B$$

The slope of the budget constraint is P_B/P_A . The slope represents how much of good A the consumer would have to give up to buy more of good B.

Graphically, the budget constraint is:



An increase in the price of good A would make the slope of the budget constraint less steep, and an increase in the price of good B would make it steeper. An increase in income would create a parallel shift of the budget constraint to the right. Suppose the consumer has an income of \$10. Good A costs \$1 per unit and good B costs \$0.50 per unit. So, the consumer could buy 10 units of A or 20 units of B. The budget constraint is:

$$\text{Income} = \text{Price}_A \times \text{Quantity}_A + \text{Price}_B \times \text{Quantity}_B$$

$$10 = 1 \times Q_A + .5 \times Q_B$$

Solving for Q_A in terms of prices yields:

$$Q_A \times P_A = I - Q_B \times P_B$$

$$Q_A = I/P_A - (Q_B \times P_B)/P_A$$

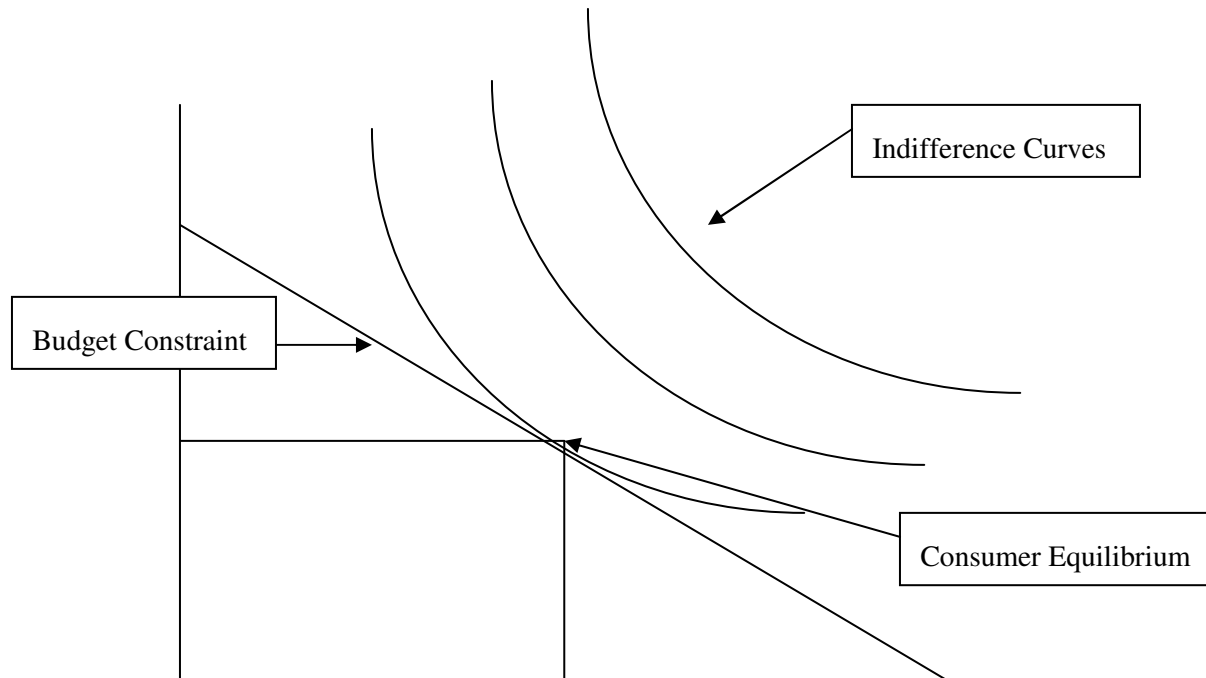
or

$$Q_A = 10/1 - .5Q_B/1 = 10 - .5Q_B$$

Consumer Equilibrium

Learning Objective: Find the equilibrium bundle of goods for a consumer, using utility analysis.

A consumer wishes to maximize utility, subject to budget or income constraints. Combining indifference curves and budget constraints:



Consumer equilibrium is at the point where a utility curve, the highest utility curve, is tangent to the budget constraint line. At this point the indifference curve and the budget constraint have the same slope. Here, the marginal rate of substitution is P_B/P_A , the rate at which the consumer must sacrifice good A for good B. Purchasing other than the equilibrium amounts of goods A and B would not be the best use of income.

©Allen Resources, Inc. All rights reserved.
Warning: Copyright violations will be prosecuted.

Any use of these materials without the express written consent of the publisher is a violation of federal and/or international copyright laws.

3. *Supply and Demand: Company Analysis (Prerequisite)*

Candidate Note: Candidates are expected to be familiar with the material covered in this prerequisite economics reading.

Learning Objectives

This summary includes a review and an analysis of the principles set forth by CFA Institute. Upon review of this summary, you should be able to:

- ❖ *Understand and calculate accounting profit, economic profit, normal profit, and economic rentpg. 38*
- ❖ *Understand, calculate, and compare total, average, and marginal revenuepg. 40*
- ❖ *List the firm's factors of productionpg. 41*
- ❖ *Understand and calculate total, average, marginal, fixed, and variable costspg. 43*
- ❖ *Understand how the profit-maximizing level of output can be determinedpg. 45*
- ❖ *Explain the differences between short-run and long-run profit maximizationpg. 45*
- ❖ *Explain decreasing cost, constant cost, and increasing cost industries and the law of supply in eachpg. 46*
- ❖ *Understand and calculate total, average, and marginal product of laborpg. 47*
- ❖ *Understand the optimal combination of resources that minimizes costspg. 51*

Overview

This summary looks at revenues, costs, and the resulting profits, including how profits and returns are measured. Finance theory says that the goal of the firm is shareholder wealth maximization. Shareholder wealth is maximized by maximizing share price which is in turn maximized by accepting all possible positive net present value projects.

Profits

Learning Objective: Understand and calculate accounting profit, economic profit, normal profit, and economic rent.

Costs serve to allocate resources in an economy such that the goods most highly valued by consumers are produced. Since the production of one good necessitates a reduction in the production of other goods, producers signal through the price, which goods they value most. The cost of production is essentially the most highly valued opportunity that was forgone in order to produce the good. The production cost of a good reveals the value of the resources used to produce it, with this value stated in terms of other opportunities relinquished.

Explicit costs are *what a firm pays to purchase productive resources*. Examples include wages, interest payments, and rental payments. These are also called accounting costs.

Implicit costs are *the opportunity costs arising from using its existing resources*. These costs do not involve a direct money payment. For example, if a firm uses a plot of land it owns to build a factory, it forgoes the revenue it might have earned if the land was used as a parking lot. The potential revenues from the parking lot project are an implicit cost charged to the factory project cash flows.

Implicit rental rates are *rental incomes foregone because the firm used its capital to fund its own projects*. The implicit rental rate includes:

- *economic depreciation* - This is the change in market value of capital during a specified period. For example, suppose a piece of property could have been sold for \$1,000 and a year later it could be sold for \$800. The \$200 difference represents economic depreciation, and is an implicit cost of using capital during the period.
- *foregone interest* - If the \$1,000 had been used for some other purpose, it would have earned a return. That foregone return is a cost of capital.

Owners' resources refer to the firm owner, who supplies entrepreneurial ability to his or her firm. Normal profit is the return an entrepreneur expects to receive, on average. It is an opportunity cost, because it represents the cost of a foregone opportunity - the opportunity to run a different firm. A firm's owner can also supply labor to the firm and receive wages. The opportunity cost of working for the firm is the wages foregone by not working at the best alternative job.

Total costs are the *implicit plus the explicit costs to the firm. This figure also includes an imputed normal rate of return for the firm's equity capital.*

Opportunity cost of equity capital is *the rate of return required by investors to supply funds for a project. This rate of return is determined by what other opportunities of equivalent risk the investor has for those funds.* If the rate of return they are earning on their investment in the firm is less than the normal market rate they could earn elsewhere, they will remove their investment.

Economic profit is *the amount of profit investors receive over and above their opportunity cost of investing.* This is also called super normal or abnormal profit. For example, if the required return on a project was 10%, and the investor received a 12% return, the additional 2% represents economic profit to the investor. Zero economic profits means the investor is receiving exactly the market rate of return on the investment in the firm. The goal of the firm is to maximize shareholders' wealth, which it accomplishes by maximizing stock price, which in turn is maximized by maximizing economic profits - accepting all possible positive net present value projects.

Normal profit is *the amount of profit in which costs are exactly covered by revenues.*

Accounting profits are *the sales revenues minus the expenses during a specified time period.* Accounting profit is normal profit plus economic profit.

Accounting costs may be different from economic costs. Accounting rules may leave out implicit costs, for example, the owner's labor services or capital. While this applies primarily to small businesses, accounting rules may also leave out the opportunity costs of alternate uses of a resource used in production. Using accounting costs may understate the actual economic costs and therefore overstate accounting profits.

To maximize economic profit, a firm must consider what and how much to produce, what production techniques to use, how to organize the firm, how to compensate employees, how to market and price its goods, and whether to produce or buy from other firms. Economic profit can be calculated by looking at revenues and subtracting the costs, including the opportunity costs of producing those revenues.

Economics: Microeconomics and Macroeconomics

Here is an example of accounting for economic profit:

	Total Cost/Revenue	
Sales		\$1,000
Opportunity Costs		
Supplies used in production	\$300	
Labor to produce goods	\$200	
Total explicit costs		-\$500
Entrepreneurial wages foregone	\$100	
Interest foregone	\$150	
Economic depreciation	\$100	
Normal profit	\$50	
Total implicit costs		-\$400
Total costs		-\$900
Economic profit		\$100

Economic rent refers to *what owners of production factors receive over the minimum amount required*. **Opportunity cost** refers to *what would be received if that factor were employed in the next best alternative use*. Economic rent versus opportunity cost depends on the elasticity of supply for a particular factor.

Marginal Revenue

Learning Objective: Understand, calculate, and compare total, average, and marginal revenue.

The demand for factors of production, like labor and capital, are **derived demand** since *demand is derived from the demand for goods and services*.

The **marginal revenue product** is *the revenue from using one more unit of labor*. As the quantity of labor supplied increases, marginal revenue product decreases. In perfect competition, marginal revenue product declines because marginal product is declining. In a monopoly or monopolistic competition, marginal revenue product declines because both marginal product and marginal revenue decrease when more labor is employed, which results in declining marginal revenue product.

Total revenue is *the price per unit times the number of units produced*. **Average revenue** is *total revenue divided by the amount sold*. **Marginal revenue** is *the added revenue from selling an additional unit*. For example, suppose a firm sells its product for \$10/unit, and it sells 100 units. Total revenue is $\$10 \times 100 = \$1,000$. Average revenue is $\$1,000/100 = \10 . Suppose the firm sells an additional unit for \$10. In this case, marginal revenue is \$10 or $\$10 \times 101 = \$1,010$, divided by 101 = \$10. If the firm sells all of its units at the same price, then marginal revenue equals average revenue. If the firm, for example, decreases price as more units are sold, then marginal revenue is less than average revenue. For example, suppose the firm sells 200 units. The first 100 units cost \$10 and for orders of more than 100 the cost is \$9/unit. In this case, total revenue = $100 \times \$10 + 100 \times \$9 = \$1,900$. Average revenue = $\$1,900/200 = \9.50 . Marginal revenue is \$9. Average revenue states the price per *all* units sold. Marginal revenue is the price for an additional unit. Often, marginal revenue is the more relevant number - a firm needs to know the revenues from an additional unit sold, which it can in turn use to compute additional revenues and profits from the additional units sold.

Note:

$$\text{Marginal Product (MP)} = \frac{\Delta \text{Total Product}}{\Delta \text{Labor Quantity}}$$

$$\text{Total Revenue (TR)} = \text{Price (P)} \times \text{Total Product (TP)}$$

$$\text{Marginal Revenue Product (MRP)} = \text{Marginal Revenue (MR)} \times \text{Marginal Product (MP)}$$

$$\text{Marginal Revenue Product (MRP)} = \frac{\Delta \text{Total Revenue}}{\Delta \text{Labor Quantity}}$$

Organizing Production

Learning Objective: List the firm's factors of production.

The factors of production are inputs the firm needs to produce revenues and include the following:

- labor from skilled and unskilled workers
- land, on which the business is located
- capital, which includes the equipment and structures
- material, including any inputs needed in the production process

A firm can organize production through:

- **command systems** - *organizing using a managerial hierarchy*
Upper level managers make decisions and pass down those decision to the workers who report to them. Information moves up, from lower level employees to management. A prime example of this is the military. Firms may also operate this way, although generally with less structure than the military. Although managers may do their best to make informed decisions, they will not know as much as the employees who are running the firm day to day. When it is easy to monitor performance, for example, on an assembly line, command systems are often used.
- **incentive systems** - *creating ways, generally through compensation, to induce employees to work to maximize firm profits*
An example is a sales person who works on commission. They have the incentive to sell as much as possible to maximize their personal earnings. This, in turn, maximizes firm revenues. When it is costly to monitor performance (for example, the performance of the CEO) incentive systems are often used.

The business firm is designed to organize raw materials, labor, and machines with the goal of producing goods and/or services. To accomplish this, firms buy productive resources, transform resources into other commodities, and sell the new product or service to consumers.

Incentives to Operate Efficiently

A small business owner has a strong incentive to operate efficiently and provide good customer service. Since the small business owner receives all of the residual profits of the firm, the lower the costs and the happier the customers, the greater the profitability of the firm and the greater the value to the small business owner. In a large firm, however, shareholders have a residual claim on the firm's assets, but managers actually run the firm.

Managers may not always act in the best interest of shareholders, for example, requiring high salaries, a company jet, and other costly perks. For a large firm, the factors that promote cost efficiency and customer service include:

- a competitive environment - A firm is continually being monitored by outside investors. These investors will buy if they think the company is well managed and profitable and will sell their stock if they believe the opposite is true. Customers also monitor the firm. If they are getting good quality products and services they will continue to patronize the firm. They will go to competing firms if they receive poor service.
- correctly structured managerial compensation - More companies are providing incentive pay, not only to top managers, but to mid and lower level employees as well. Managers who perform well, as reflected by higher stock prices or higher earnings, may receive bonuses or grants of company stock or options. When managers' personal wealth is more closely aligned with the fortunes of the company, they may be more likely to act in the best interests of shareholders.

- the market for corporate control - A poorly run company, with a depressed stock price, is more likely to be a takeover candidate than a well run company. Similarly, managers of poorly run companies are more likely to lose their jobs if the company is taken over. Managers will work to avoid this unfavorable outcome by managing the company well, satisfying stockholders and customers.

Costs

Learning Objective: Understand and calculate total, average, marginal, fixed, and variable costs.

Economists often speak of the **short run** as *a time period so limited that the business cannot alter its factors of production*, for example, the size of its factory.

The **long run** is *a time period of sufficient length to let a business alter its factors of production*. All firm resources are variable in the long run.

Total fixed cost (FC) is *the sum of the costs that are constant with output*. Examples include property taxes, depreciation, and insurance premiums.

Average fixed cost (AFC) is *the total fixed cost divided by the number of units produced*. AFC declines with rising output because the total cost is spread over a greater number of units of output.

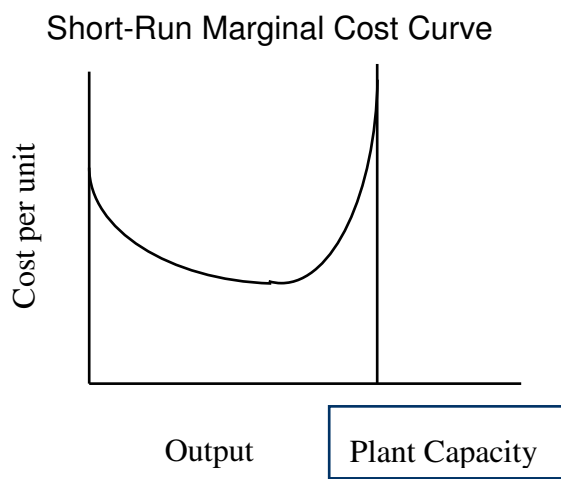
Total variable cost (VC) is *the total of those costs that rise with output*, for example labor costs.

Average variable cost (AVC) is *the total variable cost divided by the number of units produced*.

Average total cost (ATC) is *the total cost (fixed plus variable) divided by the number of units of output produced*. ATC is sometimes referred to as “per unit cost.” ATC is also equal to the sum of the average fixed and average variable costs.

Marginal cost (MC) is *the change in total cost that results from the production of one additional unit*. MC usually declines at smaller output levels and then eventually rises as output is increased.

The short-run marginal cost curve shows a sharp increase as the firm's production capacity is approached:



The average variable cost curve is U-shaped. It declines at smaller output levels and then increases as output increases. The average fixed cost curve is downward sloping. Average fixed costs are high when output is low. Average fixed costs fall as output increases. The average total cost curve is U-shaped. AFC is high when output is low, but marginal costs (and therefore total costs) will be high when the plant is nearing its production capacity.

Reasons that cost curves may shift:

- *resource prices* - As the prices of resources rise, the firm's cost curves will shift upward.
- *taxes* - The firm's average total and marginal cost curves shift upward by the amount of the tax (a component of a firm's cost).
- *technology* - A technological improvement shifts the firm's cost curves downward, reflecting the reduction in the amount of resources needed to produce alternative levels of output.

The costs a firm incurs result from the choices its managers make. Managers know what has happened in the past but face uncertainty about the future. The economic way of thinking about costs classifies them into relevant and irrelevant costs. Only relevant costs should be considered in the decision process. Opportunity costs, as discussed next, are relevant while sunk costs are economically irrelevant.

Profit Maximizing

Learning Objective: Understand how the profit-maximizing level of output can be determined.

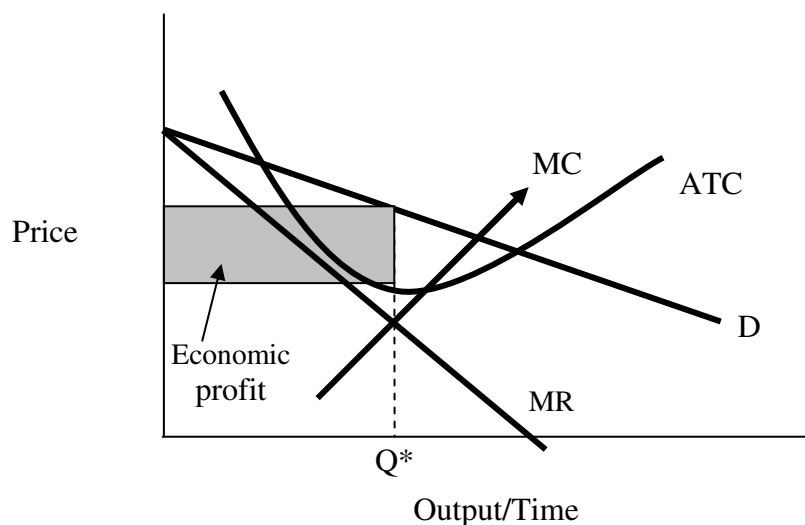
Learning Objective: Explain the differences between short-run and long-run profit maximization.

The profit-maximizing firm will *lower price and expand output until marginal revenue equals marginal cost*. Because the firm's product faces a downward sloping demand curve, a price reduction that increases sales exerts two conflicting influences on total revenue:

1. Total revenue increases because the number of units of output rises.
2. Total revenue decreases because the price is now lower on units that would have been sold at a higher price.

The second consequence of a lower price explains why the firm's marginal revenue curve lies below the demand curve. By lowering the price, the firm sells more units. The lower price is not exclusively charged on the additional units sold; rather, the lower price is now the price for all units sold. Prior to the price change, the firm could have sold some units at the higher price. Now these units are being sold at the uniformly lower price. Thus, the additional revenue associated with the sale of an additional unit of output is less than the price paid for the unit because the firm is earning less per unit on *all* units sold.

Price and Output for a Price-Searcher



Economics: Microeconomics and Macroeconomics

A firm producing at $MR = MC$ may earn an economic profit. However, under free entry into the market, rival firms will erode these economic profits. The entrance of these firms increases the supply in the market and drives down price. The price falls until the economic profit is zero and no new firms have an incentive to enter the industry. In the long run, economic profits are zero in free entry markets.

The **law of diminishing returns** dictates that *as increasing units of a variable factor are applied to a fixed amount of other resources, total product will increase, first at an increasing rate and then at a decreasing rate*. This is because the returns to the variable factor diminish as the use of the input increases beyond a certain level. The authors use as an example a farmer that applies fertilizer more and more intensively to an acre of land (the fixed factor). After a certain point, applications of more fertilizer increase the yield of wheat, but by successively smaller amounts.

A firm facing diminishing returns requires larger and larger additions of the variable factor to expand output by one unit. This causes marginal cost to rise. As MC continues to rise it eventually exceeds average total cost. As long as MC is below ATC, the contribution of MC is to bring ATC lower and lower. However, once the additional cost is greater than the average cost, ATC must start to rise.

The **Long-Run Average Total Cost Curve (LRATC)** is also called the **planning curve**. The LRATC reflects *the expected per unit cost of producing alternative rates of output while long-run factors of production are still variable. The LRATC shows the minimum average cost of producing each output level when the firm is free to choose among all possible plant sizes*.

Constant, Decreasing, and Increasing Cost Industries

Learning Objective: Explain decreasing cost, constant cost, and increasing cost industries and the law of supply in each.

In an **increasing cost industry** *prices and costs are higher when there is higher industry output in the long run*. The price increase leads to more economic profit, and over the long run, competitors will enter the market in response to the greater economic profits, increasing supply. Costs will increase, leading to a new equilibrium at both higher prices and higher supply. Consumers see this in the oil industry - when input costs and demand increase, prices also rise. The long-run supply curve is upward sloping. In a **decreasing cost industry** *supply increases leading to lower input prices*. Firms respond by lowering prices. Such efficiencies can result from new technology, economies of scale, or larger firm size. The long-run supply curve is downward sloping. In a **constant cost industry**, there are *no changes in input or output prices*. The long-run supply curve is horizontal.

Labor

Learning Objective: Understand and calculate total, average, and marginal product of labor.

Although one typically thinks of supply and demand setting the price for goods (i.e., oranges, pencils, steel, etc.), the laws of supply and demand apply to any situation in which there is scarcity; that is, any situation where something desirable exists, and the supply of which is finite. In such a case, some price will be created, even if that price is not expressed in terms of currency.

Production Terms

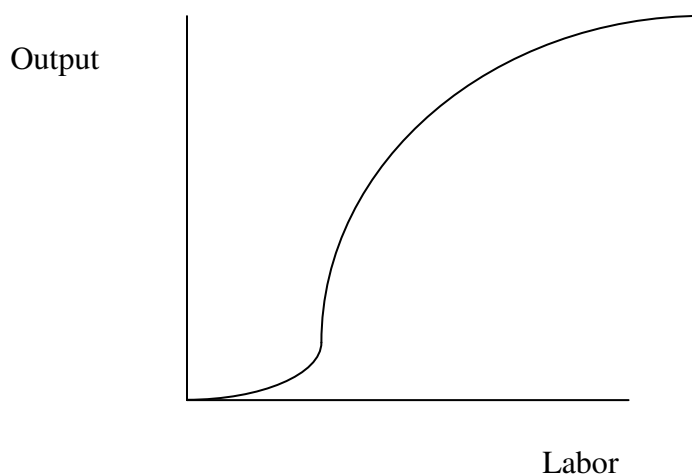
Total product is *the total output of a good that is related to a specific amount of a variable input*. Total product changes as the variable input changes.

Marginal product (MP) is *the increase in the total product due to a unit increase in the use of a variable input*. The MP curve rises at first and then declines as the use of the variable input increases.

Marginal product of labor is *the rise in total product following a one unit increase in the quantity of labor used, with all other inputs remaining the same*.

Average product (AP) is *the total product divided by the number of units of the variable input needed to produce the output level*. The AP curve rises at first and then declines as the use of the variable input increases.

Product curves *graph the relationship between employment and total, marginal, and average product*.



Economics: Microeconomics and Macroeconomics

Points above the curve are not attainable, points below the curve are attainable but inefficient, and points on the curve are attainable and efficient.

Marginal product curve: Marginal product is represented as the slope of the total product curve. The marginal curve shows how much output is added for each unit of labor added.

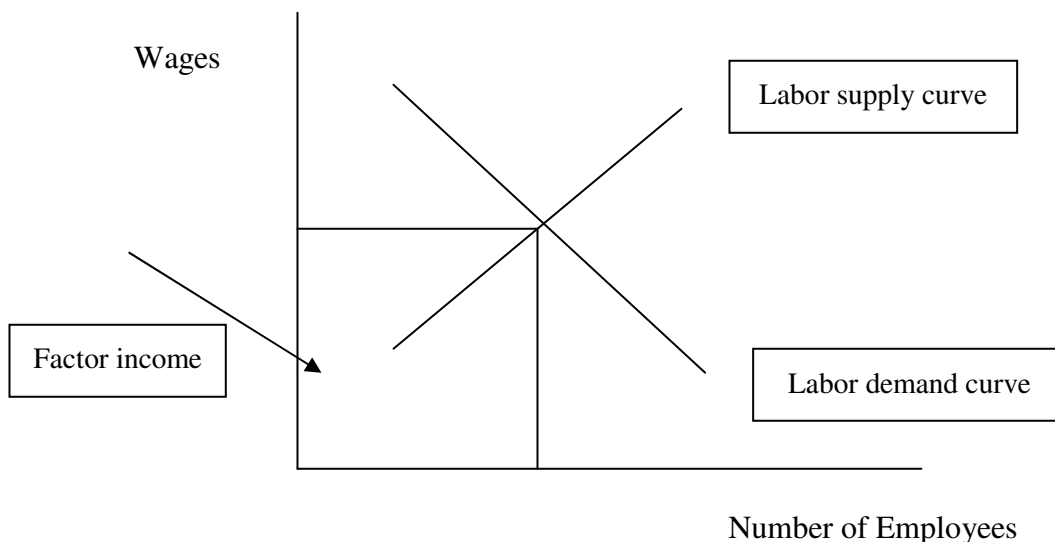
Decreasing marginal returns means that as an additional unit is produced, there is a positive return, but it is smaller than the return received from the previous unit. This could occur, for example, because the market is saturated with the units, and new units can only be sold for a lower price. Diminishing returns could also occur because at some point in the production process, using more labor or more capital does not provide a proportional investment; the process may already be as efficient as it can be.

Increasing marginal returns is the opposite: as an additional unit is produced, returns are positive and more than they were for the previous unit produced. This could occur, for example, because there is a learning curve, and as more units are produced, the production process becomes more efficient, and units are produced at less cost per unit.

The Labor Market

An employee (laborer) essentially sells his/her time, knowledge, skills, etc. to an employer. Therefore, in the labor market, laborers are *suppliers* and employers are *consumers*.

The convergence of supply and demand will set the wage for all skill sets that exist in an economy. This pricing system will then ration laborers across employers in the exact manner that prices ration goods across consumers. There is a strong relationship between the demand for a good and the demand for the resources to produce that good, including labor. Therefore, employees with the skills to produce goods in high demand are likely to command higher wages. Conversely, goods which require specialized (high-priced) labor to produce will be more costly.



When MRP equals wages and when $MR = MC$, profits are maximized. $MRP = MR \times MP$ and $MRP = W$ (wages) at the point where profits are maximized, so $W = MR \times MP$. Rearranging terms, $MR = W/MP$.

Demand for labor is affected by:

- *Output prices* - Higher prices lead to a higher demand for labor. This is because higher prices increase marginal revenue, which increases the marginal revenue product of labor. If output prices increase, then the demand for labor increases and the demand curve moves to the right.
- *Other factor prices* - In the long run, when other factors of production can be changed, the demand for labor will change. The direction of the change depends on whether the factor is a substitute for labor (more automation, fewer employees) or whether it is a complement (more equipment, more equipment operators).
- *Technology* - Technological advances create some jobs (software creators and computer programmers) and destroy others (buggy whips).

Labor Demand Elasticity

Labor demand elasticity *measures how responsive labor demanded is to wage rates*. In general, if supply increases, wages decrease. If demand is inelastic, then labor income decreases. If demand is elastic, an increase in supply means lower wages and higher labor income. If labor demand is unitary elastic, then labor income is unchanged if supply changes.

Labor demand is less elastic over the short run. This is because only quantity of labor can be changed. Over the long run, quantity of labor, as well as other factors of production, can be changed.

Elasticity of labor demand is affected by:

- *Labor intensity* - Labor-intensive industries will be more sensitive to wage rate changes. For example, for a labor-intensive industry, if wages go up significantly, then quantity of labor demand will decrease by a large amount.
- *Elasticity of product demand* - If demand for the product is elastic, then elasticity of demand for labor will be higher. If demand for a product with highly elastic demand decreases, then labor demand will have a proportionately large decrease.
- *Whether capital can be substituted for labor* - Labor demand is more elastic relative to the ease of substituting capital for labor.

Substitution effect: The substitution effect is *how much of an increase (or decrease) in wages can be attributed to the fact that the wage rate is cheaper (or more expensive), relative to leisure time*. In other words, you may prefer to golf on weekends, but if you are offered a higher enough wage, you'll work weekends and forego golf.

Income effect: The income effect can be defined as *how much of the increase (or decrease) in the amount of labor supplied can be attributed to the increase (or decrease) in the consumer's*

Economics: Microeconomics and Macroeconomics

real income and the consumption possibilities available to the consumer. Workers with higher incomes will demand more leisure, which decreases the quantity of labor supplied.

At low wage rates, when wages increase, more labor is supplied. If wages continue to increase, the quantity of labor supplied decreases.

Supply of labor is impacted by:

- the number of workers. How many adults are in the population?
- technological change. How many workers are needed?
- capital accumulation. Greater capital accumulation increases the demand for labor as people demand more goods and services.

Equilibrium in the Labor Market

Wages and employment levels have both been increasing. In general, technological change has created more jobs than it has destroyed. Incomes have risen, leading to more demand for goods and services and a need for more labor to produce these goods. The supply of labor has also increased because of advances in homemaking. People spend less time on meal preparation and cleaning and have more time to devote to the work force.

Labor markets may be dominated by employers who could lower wages below the competitive level, or by labor unions with the power to raise wages above the competitive level.

Labor unions are *formed by workers who want to increase wages or improve working conditions.* A **craft union** is *a labor union comprised of workers with the same skills, working in various industries.* An **industrial union** is *a labor union including workers in the same industry, with a variety of skills and jobs.* Many unions are members of the AFL-CIO, a combination of the American Federation of Labor and Congress of Industrial Organizations. Unions were strongest in the 1950s when about a third of the labor force was unionized. The percentage of workers in labor unions has been steadily declining since then, to around a tenth of the labor force now.

The process of negotiating with employers is **collective bargaining**. Workers may choose to strike, in other words, refuse to go to work, as a strategy to get employers to comply with their demands. Employers may lockout employees, closing down plants, rather than complying. If employers and union representatives cannot agree, they may use binding arbitration, which involves turning to a neutral third party to settle their differences.

Union goals include:

- higher wages
- better working conditions
- greater job opportunities

Labor unions face two constraints. The supply side constraint is the unions' ability to keep nonunion labor from working in the same markets as union labor. A profession where there are certifications required (electricians, plumbers) or professional associations (accountants, physicians) are better able to control entry into the work force. The demand side constraint is that as the union increases wages, there is less labor demanded.

Unions try to change the demand for labor. For example, if they can make the demand for labor less elastic, then there will be less of a drop in demand for labor if wages rise. Unions also try to increase the demand for union labor, which will have an even more beneficial impact. This could be accomplished through:

- increasing the marginal product of union labor, through, for example, training, apprenticeship, and certification programs.
- supporting higher minimum wages, which will decrease demand for low-skilled labor and increase demand for its substitute, skilled labor.
- supporting import restrictions and immigration restrictions.
- increasing demand for goods, for example, through campaigns to buy union-produced goods.

Union workers do, on average, earn more than nonunion workers, a wage gap of 10 to 25 percent for union versus non-union workers holding the same type of job.

Learning Objective: Understand the optimal combination of resources that minimizes costs.

Firm profits are limited by:

- **technology** (methods used to produce goods and services) - Whenever a firm wants to produce more, it must spend money on resources. While there are always advances in technology at any point in time, the firm's profits are constrained by the technology available.
- **information** - A firm always deals with missing information, for example, what the economy will be like in the future, the future inflation rates, future prices for raw materials needed, future customer tastes, etc. A firm must always deal with some uncertainty. Having limited information also limits profits.
- **the market** - The firm's profits are limited by the prices consumers are willing to pay for its goods, by the willingness of employees to work for the firm, and the willingness of investors to place their capital in the firm.

Production efficiency can be divided into technological and economic efficiency. Technological efficiency refers to achieving production using the least amount of inputs. A firm is operating with economic efficiency when it is producing with the least cost achievable.

Economics: Microeconomics and Macroeconomics

Technological efficiency looks at the different methods that can be used for production, and chooses the method that requires the least amount of inputs. There may be several methods of production that are technologically efficient. For example, suppose you could produce 100 units of a good with 10 units of labor and 20 units of equipment. You could produce the same 100 units using 20 units of labor and 20 units of equipment. The second method is not technologically efficient because it requires more labor and the same amount of equipment for production.

Economic efficiency requires the firm to look at the specific costs of each method that is technologically efficient and determine which method produces a given amount for the least cost. Technological efficiency looks at all available production methods, while economic efficiency looks at the relative costs of each feasible method.

Suppose you have alternate ways of producing widgets, one more labor intensive and less capital intensive and another more capital intensive and less labor intensive. The third method is equal in its use of labor and capital.

Method	Labor	Capital
1	50	100
2	100	50
3	50	50

None of these methods is technologically inefficient. More labor means less capital and visa versa.

Suppose you had a Method 4 where labor = 50 and capital = 60. This method is technologically inefficient because it uses the same amount of labor as Method 3, but more capital.

Suppose you can assign a cost to labor and capital, with labor costing \$1 a day and capital \$5.

Supply and Demand: Company Analysis (Prerequisite)

The costs for each method are as follows:

Method	Labor Cost	Capital Cost	Total Cost
1	$\$1 \times 50 = \50	$\$5 \times 100 = \500	$\$50 + \$500 = \$550$
2	$\$1 \times 100 = \100	$\$5 \times 50 = \250	$\$100 + \$250 = \$350$
3	$\$1 \times 50 = \50	$\$5 \times 50 = \250	$\$50 + \$250 = \$300$
4	$\$1 \times \$50 = \$50$	$\$5 \times 60 = \300	$\$50 + \$300 = \$350$

Method 3 is the most economically efficient because it uses the least capital, given the amount of labor needed.

this page intentionally left blank

4. *Supply and Demand Analysis*

Learning Objectives

This summary includes a review and an analysis of the principles set forth by CFA Institute. Upon review of this summary, you should be able to:

- ❖ *Understand and calculate price, income, and cross-price elasticities of demand, along with the factors that affect each of thesepg. 56*
- ❖ *Understand income and substitution effectspg. 60*
- ❖ *Explain the differences between normal and inferior goodspg. 60*
- ❖ *Explain diminishing marginal returnspg. 62*
- ❖ *Explain breakeven and shutdown points of productionpg. 63*
- ❖ *Understand the effects of economies and diseconomies of scale on costspg. 64*

Overview

Economics is broadly categorized into macroeconomics and microeconomics. Macroeconomics deals with aggregate economic quantities while microeconomics deals with markets and the decision making of individuals and firms. Microeconomics is further classified into two economic units - the consumer and the firm. The theory of the consumer deals with demand for goods and services from utility-maximizing individuals, while the theory of the firm addresses the supply of goods and services from profit-maximizing firms.

Price, Income, and Cross-Price Elasticities

Learning Objective: Understand and calculate price, income, and cross-price elasticities of demand, along with the factors that affect each of these.

Earlier prerequisite notes explained how the law of demand predicts the relationship between the price of a good and the quantity of that good being purchased. However, price is not the only thing that influences the quantity purchased. The **demand function** expands on the law of demand to attempt to predict how other factors will affect the quantity of goods purchased. In addition to price, the demand function may also capture effects of consumers' income, the price of similar goods, or even consumer tastes.

Note that the demand function may incorporate the prices of numerous distinct (but possibly similar) goods, but we are still interested in predicting the amount of a good purchased (dependent variable) given the price of that good (independent variable). Economists often distinguish between the independent variable being tested and the prices of other goods used in the demand function by labeling the dependent variable as the "**own-price**" variable. Also note that we could choose to flip the demand function, so that we are solving for price in terms of quantity purchased. This is known as the **inverse demand function**.

In earlier notes, we also established the shape of the demand curve and how it can be interpreted. To quickly recap - each point of the curve is either the highest price someone would pay for a quantity of a good, or the highest quantity that can be purchased at a given price. It can be informative to understand *how sensitive the quantity demanded (or supplied) responds to changes in any independent variable of the demand function*. This is the concept of **elasticity**. Generally speaking, elasticity is calculated as the ratio of a percent change in quantity given the percent change in some variable.

The most basic elasticity measure is the percent change in quantity per percent change in the own-price. Expressed formulaically:

$$E = \% \Delta Q / \% \Delta P$$

This is the **own-price elasticity of demand**. Since we are assuming a linear demand curve, it can be helpful to think of elasticity as the slope of the demand curve. If a product's demand is

inelastic, that means demand is not sensitive to price changes; in other words, the magnitude of the variable E (ignoring sign) has a value less than 1. If demand is *elastic*, demand is sensitive to price changes and E will have a value greater than 1. A special case of elasticity occurs when the magnitude of E is equal to 1 exactly; this is called *unit elastic*. Note that a demand curve can be elastic or inelastic depending on which point of the curve we are observing. For a simple linear demand curve, the demand is unit elastic at the midpoint of the curve. Demand is elastic above the midpoint and it is inelastic below the midpoint.

There are certain attributes that will help predict how elastic the demand of a good might be. A very important characteristic is whether there are suitable substitutes for that good: having one very similar substitute, or several acceptable substitutes, will likely lead to high elasticity, while having no substitutes will cause demand to be inelastic. Other important attributes include the portion of a household budget used on that good, how vital that good is to a consumer, and how much time is available to respond to a price change. Demand tends to be *inelastic* if the good takes up a small part of a household budget, if it is considered to be vital to the consumer, or if we are dealing with a short time window. The opposite also tends to hold, where demand is usually *elastic* if a good uses a large portion of a household budget, is viewed as discretionary, or if we are working with a longer time frame.

Finally, we may also choose to focus on **total expenditure**, which is simply the product of P and Q. In other words, it is the area of a rectangle under the demand curve formed when its base is Q and height is P. The relationship between price and total expenditure varies with elasticity. In cases where demand is elastic, price and total expenditure move in opposite directions, but if demand is inelastic, price and total expenditure move in the same direction.

Any of these variables that affect elasticity can be treated as an independent variable within the demand function. Therefore, we can calculate elasticities using any number of variables. Next, we explore a few different types.

Cross elasticity of demand - refers to how responsive demand is to changes in the price of a substitute or complement for that good or service. The formula for the cross elasticity of demand is:

$$\text{Cross elasticity of demand} = \frac{\text{Percent change in quantity demanded}}{\text{Percent change in price for a substitute or complement}}$$

For example, if consumers demand 10% more of a good and the price for a substitute of that good increases in price by 5%, then the cross elasticity of demand is 2%. Cross elasticity is positive for substitute goods and services and negative for complements. The larger the cross elasticity of demand, the more the change in demand, or the more the demand curve shifts when there is a change.

Close substitutes and complements have high cross elasticities. For example, you would expect two different brands of ice cream (substitutes) or bread and butter (complements) to have high cross elasticities. Unrelated goods (books and jeans) would be expected to have low cross elasticities.

Income elasticity of demand - measures the responsiveness of the demand for a good to a change in income. It is defined as *the percent change in the quantity of a product demanded divided by the percent change in consumer income*:

$$\frac{\text{Percent change in quantity demanded}}{\text{Percent change in income}}$$

Necessities like fuel and electricity have low income elasticity, while luxuries like air travel and private education are highly income elastic.

Factors that Affect Elasticity

A change in quantity demanded refers to movement along the demand curve for a good, generally in response to a change in price of that good. Changes in demand (or shifts in the demand curve) occur when factors like income, changes in the number of consumers in the market, changes in expectations, demographic changes, changes in consumer tastes or preferences, or changes in prices of substitute goods (factors other than price) alter the decisions of consumers.

1. *Changes in income* - The demand for most goods is positively related to income. Thus, an increase in an individual's income will cause him or her to spend more on all goods (with a few exceptions).
2. *Changes in the distribution of income* - The distribution of income among consumers changes when the number of potential consumers in the market changes or the income level of each consumer changes. The demand curve for certain products shifts when there is a change in the distribution of income.
3. *Changes in related good prices* - Related goods may be either substitutes or complements. Two goods are **substitutes** if *an increase in the price of one increases the demand for the other*. An example is the substitution of Fords for Chevrolets. Demand for a good shifts outward if the price of a substitute good increases. Two goods are **complements** if *an increase in the price of one decreases demand for the other good*. Hamburgers and hamburger rolls are an example of complementary goods. Demand for a good shifts inward if the price of a complementary good increases.
4. *Changes in market demographics* - The **demographics** of a market refer to the *number and characteristics of consumers in that market*. For example, an increase in the number of elderly people in a particular market will increase the demand for medical care and retirement housing in that market.
5. *Changes in consumer preferences* - Preferences may change because people learn about consumer choices and switch between complementary goods or alter their consumption habits. Advertising may also influence preferences (discussed below).
6. *Changes in the expected future price of a good* - If consumers expect the future price of a product to rise, their current demand for it will rise. The opposite holds if consumers expect the future price to fall.

Advertising serves to convey information about product price, quality, and availability. Advertising also serves to discipline producers through consumer confidence in a brand. A brand is essentially a product or producer’s public reputation, which is put at risk whenever products are sold. If a brand is well known, repercussions would be great if a product were to disappoint. For example, if a local hotdog vendor was found to be serving contaminated food, this would only impact his stand. However, if McDonald’s had the same problem, their worldwide reputation could be destroyed, resulting in billions of dollars lost. This illustrates why a quality brand is so important to both consumers and producers.

Substitute availability has the most impact on demand elasticity. A price rise of a good with substitutes (like fountain pens) induces customers to switch to other products (like pencils or ballpoint pens). As the number of substitutes increases, elasticity increases. For example, insulin has a very inelastic demand curve. However, the demand curve for *a given brand of insulin* might be far more elastic, as substitutes exist.

Also recall that demand for most products will be more elastic in the long run than in the short run. The direct relationship between the elasticity coefficient of demand and the length of the time period allowed for consumer adjustment is often referred to as the **second law of demand**.

Firm View of the Demand Curve

A firm may view the demand curve for its products as a potential revenue curve. Each point along the demand curve represents a total revenue figure. *The price elasticity for a firm’s product determines the degree of the firm’s pricing power*, or the ability to change price in an attempt to maximize revenue.

A demand curve may have an elastic, inelastic, and a unitary elastic portion. At some points along the demand curve, elasticity will be greater than at other points. To imagine why this might be, imagine that the price of BMWs began to increase. Initially, consumers might be willing to pay more for the car, implying inelastic demand. However, as the price continues to rise, consumers begin to choose substitutes, implying elastic demand.

Therefore, even if a producer faces what appears to be inelastic demand for its goods, this does not imply prices can rise indefinitely. The table below shows the impact of price increases or decreases on total revenue, given a certain level of elasticity.

Price Elasticity of Demand	Impact of Raising Price on Revenue	Impact of Decreasing Price on Revenue
Elastic (1 to ∞)	Decrease	Increase
Unitary (1)	None	None
Inelastic (0 to 1)	Increase	Decrease

If a firm were attempting to maximize revenue, it would continue to increase or decrease the price of its good until it reached unitary elasticity. Of course, firms try to maximize profit, not revenue, and therefore the actual price selected will be impacted by the marginal cost of production.

Substitution and Income Effects

Learning Objective: Understand income and substitution effects.

Learning Objective: Explain the differences between normal and inferior goods.

Recall that the **law of demand** explains the relationship between price and quantity. A consumer will generally purchase more of a good when prices fall, or purchase less as prices rise. There are two concepts that summarize why this occurs:

1. **Substitution effect:** As the price of a good/service declines, it induces consumers to buy more of it and less of a similar, but higher priced good. The substitution effect is *how much of an increase (or decrease) in the amount of a good consumed can be attributed to the fact that the good is cheaper (or more expensive), relative to other goods due to a reduction (or increase) in price*. For example, if a consumer regularly buys two types of cereal - A and B, then with a decrease in price of cereal A, the consumer is likely to purchase more of cereal A and less of cereal B.
2. **Income effect:** As the price of a good declines, consumers who purchase that good experience an increase in income (that is, they can now buy more units of all goods). The income effect can be defined as *how much of the increase (or decrease) in the amount of the good consumed can be attributed to the increase (or decrease) in the consumer's real income, the consumption possibilities available to the consumer*.

It is important to understand that the substitution effect is always positive; i.e., consumption of a good whose price has fallen will always increase due to the substitution effect. However, the income effect can be negative or positive since the consumption of the good whose price has fallen can either increase or decrease. This is because the decrease in the price of a commodity increases the real income of the consumer, allowing the consumer to buy more of the *same* commodity or buy *some other* commodity also.

The final effect of a decrease in the price of a commodity on its consumption will depend on the magnitude of both the substitution effect and the income effect. The substitution and income effect can also be used to explain a decrease in demand of a commodity when its price increases.

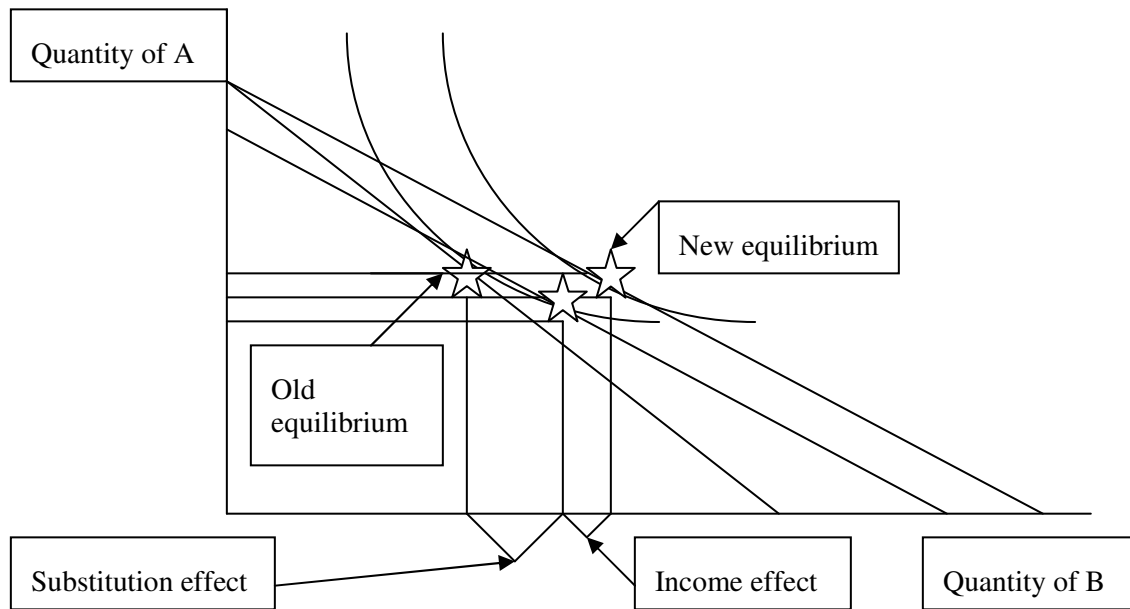
Normal goods - are *goods for which income elasticity is positive*. In other words, increase (decrease) in consumer income causes demand to rise (decline).

Inferior goods - are *goods for which the income elasticity is negative*, i.e., an increase (decrease) in consumer income causes the demand for such a good to decline (rise), generally because consumers substitute the inferior good for some other good. Most people, for instance, consider Spam (the meat product, not email) to be an inferior good; it experiences an increase in

consumer demand when income levels are lowest. We might suspect that as income rises, consumers substitute steak for Spam.

If the price of a good falls, first, it is relatively less costly when compared to other goods and it becomes a bargain. The fall in price also means real income rises; in other words, the amount of goods that can be purchased is greater, even though income itself did not change. For normal goods, this means more of the good will be purchased. In other words, both the substitution and the income effects lead to increases in purchases for a good when prices falls.

To distinguish between the two effects:



The star on the far left represents equilibrium at the old price. Assume the price of good B decreases, so more of this is purchased. The new equilibrium is the star on the far right, with less of A and more of B demanded. The indifference curve on the left is tangent to the original budget constraint and the far left star. Remember that a price decrease means an increase in real income - more can be purchased with the same income. So, assume this increase in income and the new tangency point on the same indifference curve is represented by the star in the middle. This separates the two effects into the substitution effect - the desire for B relative to A - and the income effect - the desire to purchase more because income is greater.

For an inferior good, a decrease in price means the consumer will purchase more. But, the increase in real income means the consumer wants less, not more of the good, so in this case the substitution and income effects move in opposite directions. If the income effect (less demand) is lower than the substitution effect (more demand), the net impact is that the consumer will buy more. If the income effect is greater than the substitution effect, the consumer will buy less of the good even though the price is lower. It is possible, but not likely, for the income effect to exceed the substitution effect.

This dichotomy can be illustrated with interest rates. Higher interest rates cause consumers to save more (substitution effect). Yet, higher interest rates mean one needs to save less to reach a specified goal (income effect), so it is possible for higher interest rates to result in less savings.

A **Giffen good** is *one in which the income effect is both negative and greater than the substitution effect*. By definition, all Giffen goods are inferior (but not all inferior goods are Giffen goods). This effect is named for Robert Giffen who looked at the Irish consumption of potatoes in the 19th century. Potatoes were a staple of the diet, but were an inferior good - as income rose, fewer were demanded. Because potatoes comprised such a large portion of the budget, if potato prices decreased, and income rose, then fewer potatoes would be demanded. Again, this is not stable or usual, since a firm cannot increase prices and expect demand to increase.

A **Veblen good** is *one in which price factors into desirability*. Thorstein Veblen looked at conspicuous consumption, in which a consumer gains utility by consuming high priced, status goods. Consumers might pay more for cars or jewelry because they would then be observed using these high priced goods. When the price of a Veblen good is increased, demand is higher, not lower. This creates a positively sloped demand curve. However, there is a limit to the price increase - at some point, no matter how high the status or desirability of the good, consumers will stop purchasing it, and the demand curve will become negatively sloped.

Diminishing Marginal Returns

Learning Objective: Explain diminishing marginal returns.

As explained in the prerequisite notes, firms wish to boost their marginal product by adding inputs (such as labor), but at some point the firm's output reaches a cap, beyond which any additional inputs will not boost output. **Marginal product** refers to *the improvement in productivity realized by adding one additional unit of input*. Initially, marginal product will increase, but eventually output reaches a point where the benefit from each successive unit of input will decline and ultimately reach zero benefit.

Increasing marginal returns occur when the marginal product increases when additional units of that input are deployed. **Diminishing marginal returns** occur when additional resources beyond a certain point lower marginal product, to the point where the next increase creates negative returns. Typically, a firm increases marginal returns, until a point where returns start to diminish. For example, if the firm has too few workers, adding more labor can make the workers more efficient and allow them to specialize. At some point, too many workers are added, and no additional output is produced. This occurs because usually another resource, like building and equipment, is fixed. In this case, adding more workers without adding more equipment will not increase production.

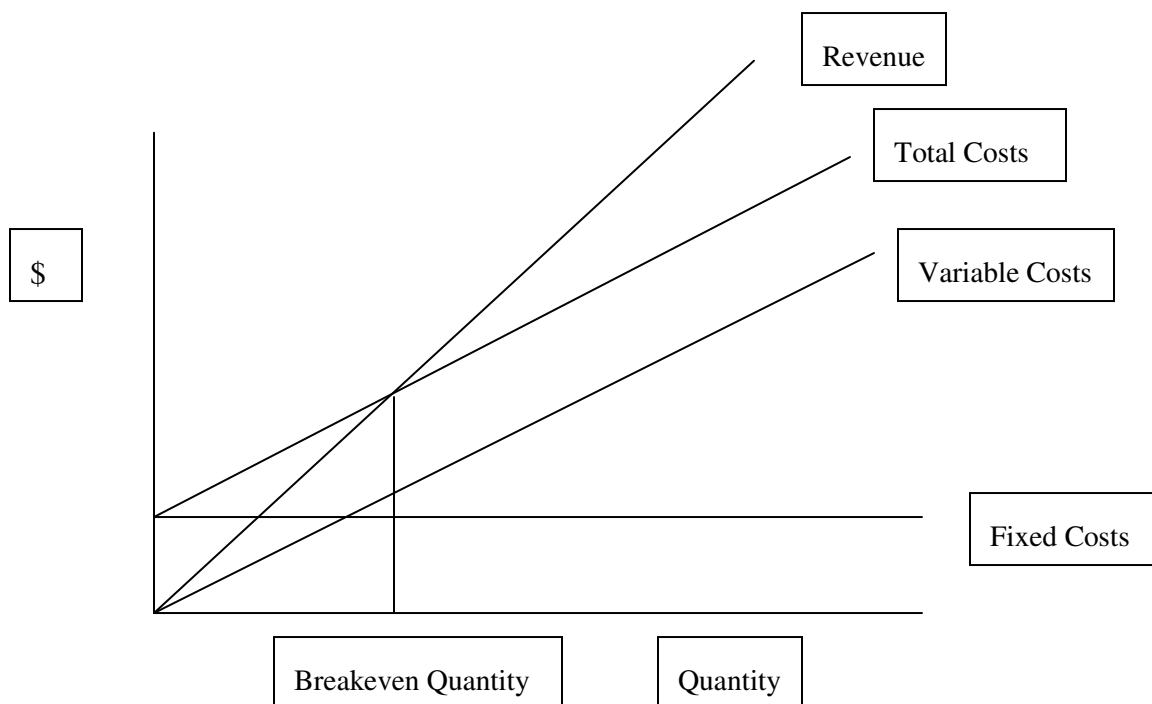
This is also true if not all workers are equally skilled. The firm will want to deploy its most efficient and productive workers first. If demand increases, the firm might need to hire workers who are not as efficient because the most skilled workers are already employed.

In the short run, with constant input prices, increasing (decreasing) marginal return decreases (increases) marginal costs and increases (decreases) profits.

Breakeven and Shutdown Points

Learning Objective: Explain breakeven and shutdown points of production.

The **breakeven point** for production is *the point at which total revenues exactly equal total costs*. Remember that total costs are the sum of variable and fixed costs. The **shutdown point** is *where average revenue is less than average variable costs*. In this case, the firm will still incur fixed costs, but will not operate if its variable costs are not covered.



There are several concepts that need to be covered to fully understand breakeven and shutdown points. First, breakeven incorporates total costs, which must capture **opportunity costs**. Any economic decision incurs an opportunity cost; in other words, when a firm decides to hire workers or use equipment, it is also choosing to forego the potential benefit from a different decision.

Therefore, it is impossible to have zero opportunity costs. A firm instead will seek to minimize opportunity cost as much as possible.

Economics: Microeconomics and Macroeconomics

Second, a firm will try to operate in a manner that will maximize its profit. There are several ways to measure profit, and economists and accountants measure profit differently. Both will start with *total revenue*. **Economic profit** is *total revenue (TR) less economic cost*. **Accounting profit** is *TR less accounting costs*. The difference is that economic costs capture opportunity cost while accounting costs do not.

Total costs can also be roughly classified as either fixed or variable, such that total fixed costs plus total variable costs equal total costs (TC). Again, we are discussing economic costs here. If $TR = TC$, a company is realizing zero economic profit (though it is likely realizing some accounting profit). This is the breakeven point. Above the breakeven point, the company is making economic profits. However, note that because of diminishing marginal returns, there will be some point at which profit will be maximized, and any additional production beyond this maximum point will cause economic profit to diminish, until a second breakeven point is reached. Increasing production beyond the second breakeven point will lead to economic losses, as variable costs will at this point outpace marginal revenues.

A company could for some period of time operate such that it is at least covering its total variable costs (TVC). A firm will shut down, however, if $TR < TVC$. If a firm's revenues can at least cover variable costs, then in the short run the firm is better off continuing operations. Why not shut down if the company is simply failing to break even? Because there will be a price range where a firm can cover all variable costs and still cover a portion of unavoidable fixed costs. In that range, losses from shutting down would exceed losses while continuing operations. As long as the situation is temporary, a company can continue to operate and eventually return to profitability. However, if the losses persist over the long run, then the company should shut down because continuing to operate would only erode any potential shareholder residuals.

The following table summarizes the decisions:

	Short Run	Long Run
$TR = TC$	Continue operating	Continue operating
$TR < TC$ and $TR = TVC$	Continue operating	Shut down
$TR < TVC$	Shut down	Shut down

Economies and Diseconomies of Scale

Learning Objective: Understand the effects of economies and diseconomies of scale on costs.

Economies of scale are *per unit cost reductions associated with using large plants to produce a large volume of output*. A firm with economies of scale will exhibit a falling long-run average total cost, or LRATC.

Diseconomies of scale *imply that long-run average total costs are greater for larger firms than for smaller ones.* Diseconomies may result from bureaucratic inefficiencies associated with larger firms. Generally, diseconomies of scale result in rising long-run average total costs.

Economies of scale arise when a firm increases all inputs so as to increase total output, and the cost per unit falls as output rises. There are several major reasons why unit costs tend to initially decline as firms produce a larger total volume of output (and use a larger size plant):

- *Adoption of mass production techniques* - Machines designed to make five times as many welds on the assembly line per day (and five times as many over their lifetime) will probably not be five times as costly. This results in increasing returns to scale, where a process allows for increases in output proportionately larger than the input increases.
- *Better equipment* - Firms can afford better/more efficient equipment.
- *Specialization* - Provides the opportunity for people to become exceptionally proficient at performing small, but essential, functions.
- *Waste reduction* - Firms can reduce waste, lower energy consumption, and improve quality control.
- *Learning by doing* - Workers and management improve their skills as they “practice” productive techniques, and they make better use of information.
- *Bulk pricing* - Firms can get discounts on inputs when purchasing in large quantities.

Diseconomies of scale arise as a result of:

- decreasing returns to scale, where increases in output are proportionately smaller than input increases.
- poor management resulting from an operation that is too large or bloated.
- duplication of functions or products.
- supply constraints that result in higher input prices.

Economic theory predicts that, at least initially, larger firms have lower unit costs than comparable smaller firms.

Economies and diseconomies of scale are long-run concepts, while increasing and diminishing returns are short-run concepts. Over the long run, competition will force firms to operate at a point where the long-run market price will be established. Therefore, the more successful firms will be those that can minimize their LRATC. The *optimal point where a firm maximizes the difference between long-run TR and LRATC*, which is in turn the point where LRATC is lowest, is called the **minimum efficient scale**. This determines the optimal firm size.

Economics: Microeconomics and Macroeconomics

In most industries, economies of scale are available at first. Constant returns to scale generally exist for a wide range of outputs. **Constant returns to scale** suggests that *unit costs remain constant when the firm's scale changes*. Neither economies nor diseconomies of scale are present. This situation is consistent with real-world conditions in many industries.

©Allen Resources, Inc. All rights reserved.
Warning: Copyright violations will be prosecuted.

Any use of these materials without the express written consent of the publisher is a violation of federal and/or international copyright laws.
