CFA® Exam Review

CFA® EXAM REVIEW

CFA PREREQUISITE ECONOMICS READINGS

ECONOMICS: MICROECONOMICS AND MACROECONOMICS

WILEY

READING 13: DEMAND AND SUPPLY ANALYSIS: INTRODUCTION

LESSON 1: DEMAND AND SUPPLY ANALYSIS: BASIC PRINCIPLES AND CONCEPTS

LOS 13a: Distinguish among types of markets.

Types of Markets

Markets may be classified into the following types:

- Factor markets are markets for factors of production (e.g., land, labor, capital). In the factor market, **firms** purchase the services of factors of production (e.g., labor) from households and transform those services into intermediate and final goods and services.
- Goods markets are markets for the output produced by firms (e.g., legal and medical services) using the services of factors of production. In the goods market, **households and firms** act as buyers.
    - Intermediate goods and services are used as inputs to produce other goods and services.
    - Final goods and services are goods in the final form purchased by households.
- Capital markets are markets for long-term financial capital (e.g., debt and equity). Firms use capital markets to raise funds for investing in their businesses. Household savings are the primary source of these funds.

Generally speaking, market interactions are voluntary. Firms offer their products for sale if they believe that they will fetch a price that exceeds costs of production. Households purchase goods and services if they believe that the utility that they will derive from the good or service exceeds the payment required to obtain it. Therefore, whenever the perceived value of a good exceeds the cost of producing it, there is potential for a trade that would make both the buyer and the seller better off.

LOS 13b: Explain the principles of demand and supply.

LOS 13c: Describe causes of shifts in and movements along demand and supply curves.

LOS 13d: Describe the process of aggregating demand and supply curves.

LOS 13e: Describe the concept of equilibrium (partial and general), and mechanisms by which markets achieve equilibrium.

LOS 13g: Calculate and interpret individual and aggregate demand, and inverse demand and supply functions, and interpret individual and aggregate demand and supply curves.

The Demand Function and the Demand Curve

Demand is defined as the willingness and ability of consumers to purchase a given amount of a good or a service at a particular price. The quantity that consumers are willing to purchase depends on several factors, the most important being the product's own-price. The law of demand states that as the price of a product increases (decreases), consumers will be willing and able to purchase less (more) of it (i.e., price and quantity demanded are inversely related). Other factors that influence the ability and willingness of consumers to purchase a good include income levels, tastes, and preferences, and prices and availability of substitutes and complements. The demand function captures the effect of all these factors on demand for a good.

Note that the law of demand need not hold in all circumstances.

$$\boxed{\text{Demand function: } QD_x = f(P_x, I, P_y, \ldots)} \quad \text{... (Equation 1)}$$

Equation 1 is read as "the quantity demanded of Good X (QD_X) depends on the price of Good X (P_X), consumers' incomes (I), and the price of Good Y (P_Y), etc."

Economists use the term own-price when referring to the price of the good that is the focus of analysis. In Equation 2, P_G (price of gasoline) represents own-price.

A hypothetical example of a demand function is the following equation, which links the per-household quantity of gasoline demanded per week (in gallons), $\mathbf{QD_G}$, to the price (in terms of dollars per gallon) of gasoline, $\mathbf{P_G}$, per-household annual income (in thousands of dollars), **I**, and the average price of an automobile (in thousands of dollars), $\mathbf{P_A}$.

$$\boxed{\text{Demand equation: } QD_G = 7.5 - 0.5P_G + 0.1I - 0.05P_A} \quad \text{... (Equation 2)}$$

We shall learn about cross-price elasticity later in the reading. Complements are goods with negative cross-price elasticity, while substitutes exhibit positive cross-price elasticity.

From the demand equation, notice that:

- The sign on the coefficient of gasoline price is negative. An increase (decrease) in the price of gasoline results in a decrease (increase) in quantity demanded. Note that this relationship conforms to the law of demand.
- The sign on consumers' income is positive. An increase (decrease) in income results in an increase (decrease) in demand for gasoline.
- The sign on the price of automobiles is negative. An increase (decrease) in the price of automobiles results in a decrease (increase) in demand for gasoline. This suggests that gasoline and automobiles are complements.

The Latin phrase "ceteris paribus" is used widely in economics textbooks. It literally stands for "all other things being equal." In this reading, we will use the phrase "holding all other things constant" to stand for the same thing.

Notice that we have used three independent variables in our example (own-price, consumers' income, and the price of automobiles). Economists typically concentrate on the relationship between quantity demanded and the product's own-price (which makes it easier to represent the relationship on a two-dimensional graph) and assume that all other independent variables that impact demand are constant when expressing the demand equation.

Let us now assume that the values of consumers' income (I) and the price of automobiles (P_A) are constant at 60 (or $60,000 per year) and 30 (or $30,000) respectively. Inserting these values in our demand equation allows us to express the relationship between the quantity of gasoline demanded and gasoline prices as:

Note that income and the price of automobiles are not ignored in this equation. They are only assumed constant and their impact on demand for gasoline is incorporated in the new constant term (12).

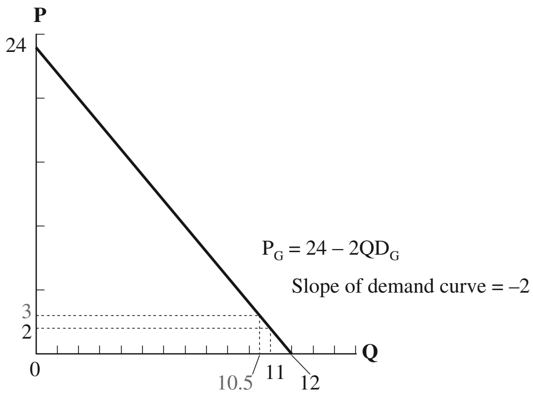
$$\boxed{QD_G = 7.5 - 0.5P_G + 0.1(60) - 0.05(30) = 12 - 0.5P_G} \quad \text{... (Equation 3)}$$

Note that Equation 3 presents quantity demanded as the dependent variable and price as the independent variable. However, economists prefer to present demand curves with quantity on the x-axis and price on the y-axis. To come up with an equation in line with these preferences, we need to invert the demand function, which basically requires us to make price the subject of the demand equation. The inverse demand function in our example is determined as follows:

$$\boxed{QD_G = 12 - 0.5P_G \Rightarrow P_G = 24 - 2QD_G} \quad \text{... (Equation 4)}$$

Figure 1-1 presents the graph of our inverse demand function, which is called the demand curve. Note that we need to restrict the value of QD_G in Equation 4 to 12 so that price is not negative.

Figure 1-1: Demand Curve for Gasoline



Note the following regarding the demand curve:

- The demand curve is drawn with quantity on the x-axis and price on the y-axis.
- The demand curve shows the maximum quantity of gasoline demanded at every given price (e.g., at a price of \$2/gallon, an individual household would be willing and able to buy 11 gallons of gasoline every week).
- Alternatively, one could also interpret the demand curve as showing the highest price a household is willing and able to pay for every given quantity of gasoline (e.g., the highest price a household would be willing and able to pay for 11 gallons of gasoline every week is \$2/gallon).
- If the price of gasoline were to rise by \$1/gallon to \$3/gallon, the quantity of gasoline a household would be able and willing to purchase would fall to 10.5 gallons per week.
- The slope of the demand curve is calculated as the change in price divided by the change in quantity demanded ($\Delta P/\Delta QD$). For our demand curve, the slope equals $(\$3 - \$2)/(10.5 - 11) = -2$.

It is very important for us to understand that the demand curve is drawn up based on the inverse demand function (which makes price the subject); not the demand function (which makes quantity demanded the subject). The slope of the demand curve is therefore not the coefficient on own-price (P_G) in the demand function; instead it equals the coefficient on quantity demanded (QD_G) in the inverse-demand function. Note that the slope of the demand curve is also the reciprocal of the coefficient on own-price (P_G) in the demand function ($1/-0.5 = -2$).

Changes in Demand Versus Movements Along the Demand Curve

It is also very important for us to understand the difference between changes in demand (shifts in the demand curve) and changes in quantity demanded (movements along the demand curve). When own-price changes (independent variable in Equation 3) there is a **movement along the demand curve** or a change in quantity demanded. When there is a change in anything else that affects demand (i.e., a change in any of the factors assumed constant in the demand function and accounted for in the intercept term in Equation 3, including consumers' incomes and the average price of automobiles) there is a **shift in the demand curve** (because the intercept term in the inverse-demand function changes) or a change in demand.

Note that we continue to assume that the price of automobiles remains $30,000.

For example, if per-household annual income were to fall to 50 ($50,000 per year versus the $60,000 per year assumed earlier when deriving Equation 3), the demand function changes to:

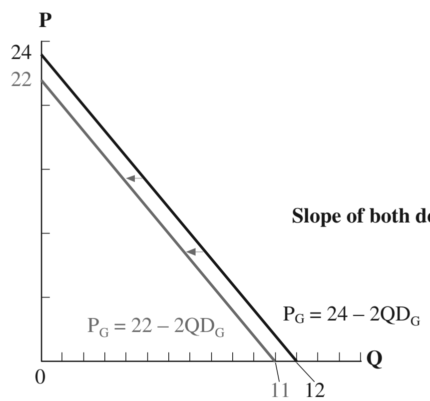
$$QD_G = 7.5 - 0.5P_G + 0.1(50) - 0.05(30) = 11 - 0.5P_G$$

The new inverse demand function (assuming that household income has fallen to $50,000 per year) is:

$$P_G = 22 - 2QD_G$$

Notice that only the intercepts have changed (relative to Equation 3). The slope has remained unchanged (slope = –2) and the demand curve has shifted inward (see Figure 1-2).

Figure 1-2: Demand Curve Before and After the Change in Consumers' Income



The shift in the demand curve illustrated in Figure 1-2 may be looked upon as a vertical shift downward or a horizontal shift to the left. For a given quantity, households are now willing and able to pay less; and at a given price, households are now willing and able to buy less. The relation between quantity demanded and own-price (slope of the inverse demand function) has remained the same.

The Supply Function and the Supply Curve

Supply refers to the willingness and ability of producers to sell a good or a service at a given price. Generally speaking, producers are willing to supply their output as long as the price is at least equal to the cost of producing an additional unit of output (known as marginal cost). The greater the (positive) difference between price and the cost of producing an additional unit, the greater the willingness to supply. Therefore, the law of supply states that price and quantity supplied are positively related.

Equations 1 and 5 are known as behavioral equations as they model the behavior of buyers and sellers respectively.

We will look at the components of costs of production (variable and fixed costs) and marginal cost in detail in a later reading. For now, we assume that the only factor of production required to produce gasoline is labor. The price of labor (the wage rate), W, is measured in dollars per hour. Therefore, the supply function can be expressed as:

$$\text{Supply function: } QS_x = f(P_x, W, \ldots) \quad \ldots \text{ (Equation 5)}$$

Equation 5 is read as "the quantity supplied of Good X (QS_X) depends on the price of Good X (P_X), and the wage rate paid to labor (W), etc."

An individual seller's hypothetical supply function for gasoline is presented below:

$$\boxed{QS_G = -150 + 200P_G - 10W} \quad \text{... (Equation 6)}$$

This function tells us that:

- For every $1 increase in the price of gasoline (own-price), quantity supplied would increase by 200 gallons of gas per week.
- For every $1/hour increase in the wage rate, the producer would be willing and able to sell 10 fewer gallons of gas per week (due to an increase in marginal cost).

In order to express quantity supplied only as a function of own-price, we assume that the wage rate is constant at $20 per hour, and come up with the following supply function:

$$\boxed{QS_G = -150 + 200P_G - 10(20) \Rightarrow QS_G = -350 + 200P_G} \quad \text{... (Equation 7)}$$

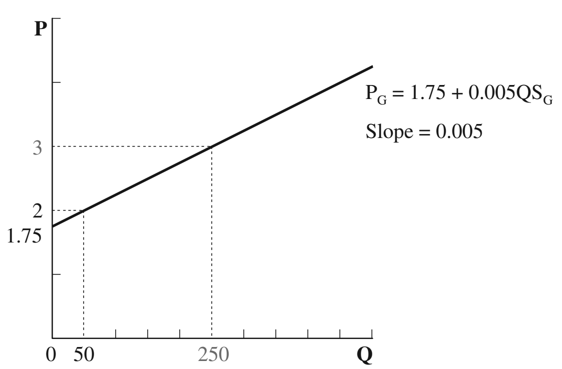
Just like the demand curve is based on the inverse demand function, the supply curve is based on the inverse supply function (with own-price as the subject). The inverse supply function in our example is given as:

$$\boxed{QS_G = -350 + 200P_G \Rightarrow P_G = 1.75 + 0.005QS_G} \quad \text{... (Equation 8)}$$

Figure 1-3 presents an individual producer's supply curve. The supply curve shows the highest quantity the seller is willing and able to supply at each price, and the lowest price at which the seller is willing and able to supply each quantity.

- At a price of $2/gallon, the producer would be willing and able to supply 50 gallons per week.
- Alternatively, the lowest price the producer would accept for supplying 50 gallons per week is $2/gallon. Finally, if prices were to rise to $3/gallon, the producer would be willing and able to supply 250 gallons per week.

Figure 1-3: Individual Seller's Supply Curve for Gasoline



It is very important for us to understand that the supply curve is drawn up based on the **inverse supply function**. The slope of the supply curve is therefore not the coefficient on own-price (P_G) in the supply function; instead it equals the coefficient on quantity supplied (QS_G) in the inverse-supply function. Note that the slope of the supply curve is also the reciprocal of the coefficient on own-price (P_G) in the supply function (1/200 = 0.005).

Changes in Supply Versus Movements Along the Supply Curve

It is also very important for us to understand the difference between changes in supply (shifts in the supply curve) and changes in quantity supplied (movements along the supply curve). When own-price changes there is a **movement along the supply curve** or a change in quantity supplied. When there is a change in anything else that affects supply (i.e., a change in any of the factors assumed constant in the supply function and accounted for in the intercept term in Equation 7, including the wage rate) there is a **shift in the supply curve** (as the intercept term in the inverse-supply function changes) or a change in supply.

For example, if the wage rate were to fall to $15/hour (from $20/hour assumed earlier when deriving Equation 7), the supply function would change to:

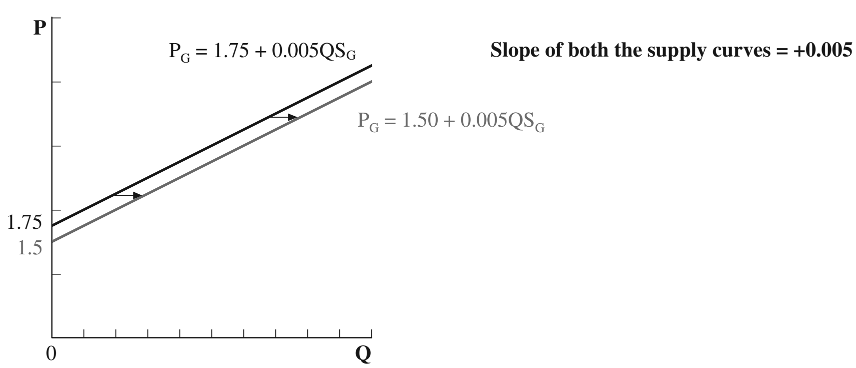
$$\boxed{QS_G = -150 + 200P_G - 10(15) \Rightarrow QS_G = -300 + 200P_G} \quad \text{... (Equation 9)}$$

The new inverse supply function is:

$$\boxed{QS_G = -300 + 200P_G \Rightarrow P_G = 1.5 + 0.005QS_G} \quad \text{... (Equation 10)}$$

Notice that only the intercept has changed (from 1.75 to 1.50), the slope has remained unchanged (slope = 0.005) and the supply curve has shifted downward and to the right (see Figure 1-4).

Figure 1-4: Supply Curve for Gasoline Before and After the Change in Wage Rates



The shift in the supply curve illustrated in Figure 1-4 may be looked upon as a vertical shift downward or a horizontal shift to the right. The producer is now willing and able to supply any given quantity at a lower price. Alternatively, at any given price, the producer is willing and able to supply more. The increase in supply has been brought about by a fall in marginal cost of production due to a decline in the wage rate. The relation between quantity supplied and own-price (slope of the inverse supply function) has remained the same.

Aggregating the Demand and Supply Functions

The demand and supply curves that we have been working with in this reading so far have represented the individual household and the individual supplier respectively. The market for gasoline however, consists of numerous consumers and producers. Individual demand and supply curves are therefore translated into the market demand and supply curves. A simple horizontal summation enables us to do this.

Aggregating the Demand Function

Let's start with the individual household demand function (Equation 2). Assuming that there are 1,000 identical households in the market, the market demand function is determined by aggregating the 1,000 individual household demand functions, or simply multiplying each household's quantity demanded by 1,000.

Market Quantity Demanded: $QD_G = 1{,}000 \times (7.5 - 0.5P_G + 0.1I - 0.05P_A)$

$$\boxed{QD_G = 7{,}500 - 500P_G + 100I - 50P_A} \quad \text{... (Equation 11)}$$

Assuming I and P_A remain constant at 60 and 30 respectively, we can express market quantity demanded of gasoline solely as a function of own-price.

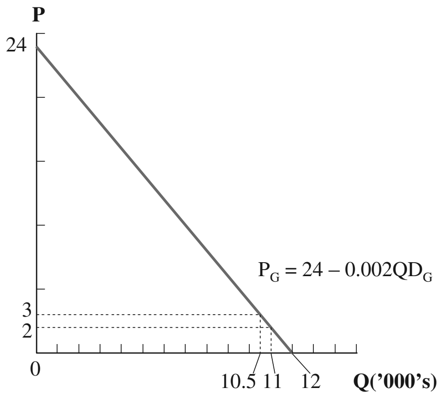
$$\boxed{QD_G = 1{,}000(12 - 0.5P_G) = 12{,}000 - 500P_G} \quad \text{... (Equation 12)}$$

The market inverse demand function is determined by making own-price the subject of the market demand function:

$$\boxed{P_G = 24 - 0.002QD_G} \quad \text{... (Equation 13)}$$

The market demand curve is the graphical representation of the **market inverse demand function** as shown in Figure 1-5. Note that we multiplied the individual demand function, not the individual inverse demand function, by the number of households to aggregate the demand function. This is because the aggregation process requires us to add up the quantities that individual households are willing and able to purchase (i.e., perform the summation horizontally), not adding their prices (which would require performing the summation vertically). If each household is willing and able to buy 11 gallons at a price of $2/gallon, then 1,000 such households would be willing and able to buy a total of 11,000 gallons at a price of $2/gallon.

Figure 1-5: Market Demand for Gasoline



Aggregating the Supply Function

To aggregate suppliers in the market, we follow the same process. Assuming 20 identical suppliers in the market, the market supply function and the market inverse supply function are given as:

$$\boxed{\text{Market Quantity Supplied: } QS_G = 20(-350 + 200P_G) = -7{,}000 + 4{,}000P_G}$$

... (Equation 14)

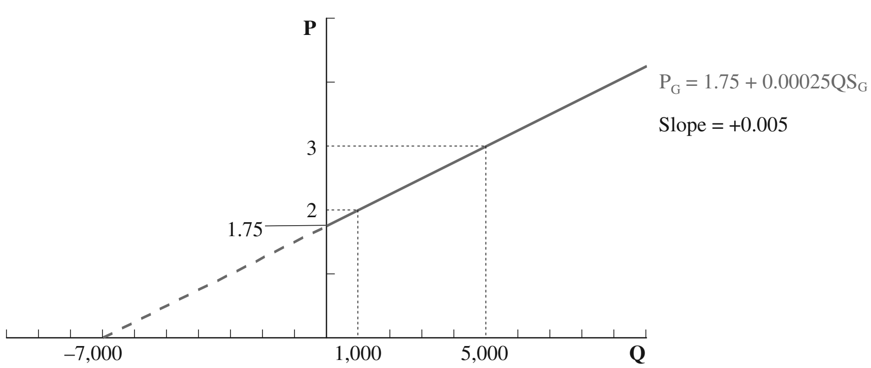
We have assumed that wage rates are constant at $20/hour. This is the same wage rate we used when deriving the individual supply and individual inverse supply functions (Equations 7 and 8).

$$\boxed{\text{Market Inverse Supply Function: } P_G = 1.75 + 0.00025QS_G}$$

... (Equation 15)

If one producer was willing and able to supply 50 gallons at a price of $2/gallon, it follows that 20 such producers would be willing and able to supply 1,000 gallons at a price of $2/gallon. The market supply curve is therefore, a horizontal summation of the individual supply curves. Figure 1-6 presents the market supply curve for gasoline.

Figure 1-6: Market's Supply Curve for Gasoline



Finally, note that the slopes of the market demand (–0.002) and market supply curves (0.00025) equal the coefficients on QD_G and QS_G in the market inverse demand function and market inverse supply function respectively.

Market Equilibrium

Market equilibrium can be defined in two ways:

1. It occurs at the price at which quantity demanded equals quantity supplied.
2. It occurs at the quantity at which the highest price a buyer is willing and able to pay equals the lowest price that a producer is willing and able to accept.

More simply, market equilibrium occurs at the point of intersection between the market demand and supply curves. We can determine the point of equilibrium by either equating the market demand function to the market supply function, or by equating the market inverse demand function to the market inverse supply function.

Equating Equations 12 to 14 (market demand function to the market supply function)

$$\boxed{12{,}000 - 500P_G = -7{,}000 + 4{,}000P_G \Rightarrow P_G = 4.22}$$

The equilibrium values of P_G and Q_G have been rounded off.

Equating Equations 13 to 15 (market inverse demand function to the market inverse supply function)

$$\boxed{24 - 0.002QD_G = 1.75 + 0.00025QS_G \Rightarrow Q_G = 9{,}888.89}$$

At equilibrium: $QD_G = QS_G = Q_G$

Therefore, given the assumed values of I (60), P_A (30), and W (20), the unique combination of quantity and price at which the gasoline market is in equilibrium is (9,888.89, 4.22).

Note that when we derived the market demand and supply functions, and the market demand and supply inverse functions, we assumed incomes (I), average prices of automobiles (P_A), and wages (W) constant at 60, 30, and 20 respectively. These variables are known as exogenous variables because their values are determined outside of demand and supply for the market being studied (gasoline). For example, the wage rate is determined in the labor market, which is different from the market for gasoline. Only own-price (P_G) and quantity (Q_G) are determined in the market being studied, so they are known as endogenous variables.

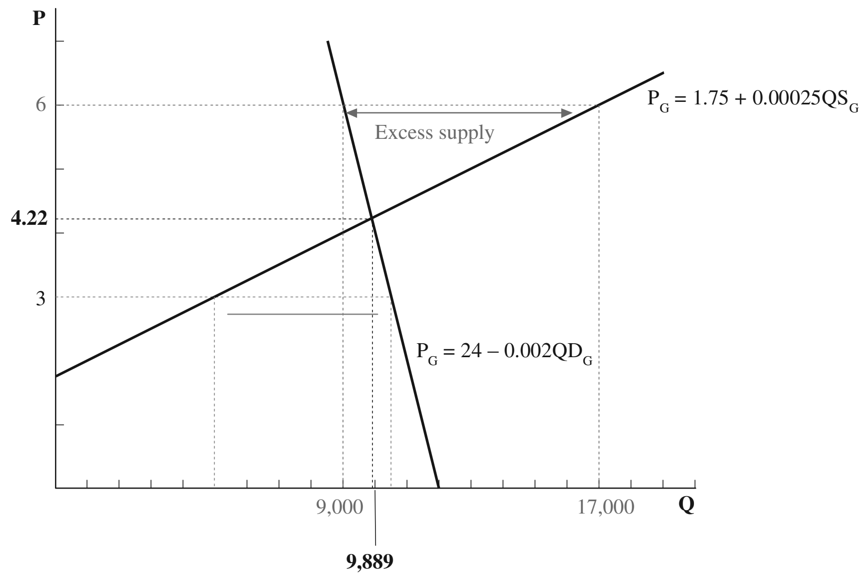
When we concentrate on one market and assume that the values of all exogenous variables are given, we are undertaking partial equilibrium analysis. Partial equilibrium analysis does not account for any feedback effects associated with other markets that are related to the one being studied. For example, if the wage rate were to rise, we would expect an increase in gasoline demand, but we have ignored the feedback effects of the wage rate to and from the gasoline market. This works while we are analyzing a very local gasoline market, but when we analyze the national market for gasoline, general equilibrium analysis, which accounts for all feedback effects to and from tangential markets simultaneously, would be more appropriate.

LESSON 2: MARKET EQUILIBRIUM

LOS 13h: Calculate and interpret the amount of excess demand or excess supply associated with a non-equilibrium price.

Iterating Toward Equilibrium: The Market Mechanism (See Figure 2-1.)

Figure 2-1: Excess Demand and Excess Supply



Suppose that in our hypothetical, the current market price is actually \$6 (above the equilibrium price of **\$4.22**). In this case, quantity demanded and quantity supplied would equal 12,000 – 500(6) = 9,000 and –7,000 + 4,000(6) = 17,000 respectively. Since quantity supplied is greater than quantity demanded at this price, there is excess supply of gasoline (of 8,000 gallons/week).

Alternatively, if the current market price is actually \$3 (below the equilibrium price of **\$4.22**), quantity demanded and quantity supplied would equal 12,000 – 500(3) = 10,500 and –7,000 + 4,000(3) = 5,000 respectively. Since quantity demanded is greater than quantity supplied at this price, there is a shortfall (excess demand) of gasoline (of 5,500 gallons/week).

LOS 13f: Distinguish between stable and unstable equilibria, including price bubbles, and identify instances of such equilibria.

If the demand curve is downward sloping and the supply curve is upward sloping (as is the case in our hypothetical), the market mechanism will always result in stable equilibrium.

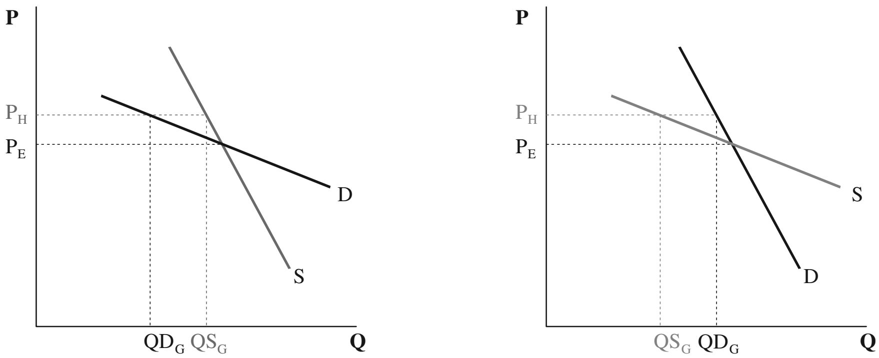
In both the above scenarios, once market prices have been bumped away (for whatever reason) from equilibrium levels, the market mechanism would direct the market back **toward equilibrium** over time and to then stay there. Such equilibrium is known as stable equilibrium. When there is excess supply, prices will fall, and when there is a shortfall, prices will rise. If there is neither excess supply nor a shortfall, the market will remain in equilibrium.

However, sometimes the market mechanism may continue to drag the market **away from equilibrium**. Such equilibrium is known as unstable equilibrium. To illustrate this, let's assume that we are working in market conditions where demand and supply are both downward sloping. We work with two scenarios, one where supply intersects demand from above (see Figure 2-2a) and the other where supply intersects demand from below (see Figure 2-2b).

For example, the supply of labor is downward sloping at relatively high wage rates where an increase in the wage rate causes workers to supply fewer hours of labor as they attach more value (utility) to an extra hour of leisure than they do to the extra hour of wages.

Figure 2-2: Stability of Equilibria

2-2a: Stable Equilibrium **2-2b: Unstable Equilibrium**



Supply intersects demand from above. At P_H: $QS_G > QD_G$ → Excess supply. P↓ to P_E so equilibrium is dynamically stable.

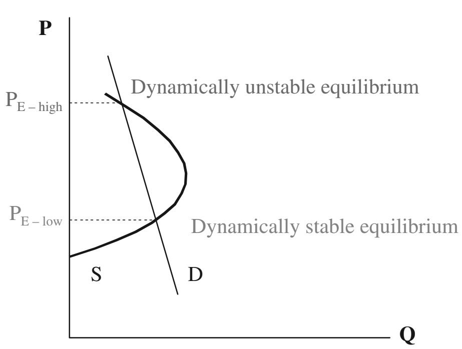
Supply intersects demand from below. At P_H: $QS_G < QD_G$ → Excess demand. P↑ away from P_E so equilibrium is dynamically unstable.

In Figure 2-2a, the supply curve is steeper than the demand curve, and intersects the demand curve from above. At price levels greater than equilibrium price (P_E), for example at P_H, quantity supplied (QS_G) exceeds quantity demanded (QD_G) so there is excess supply. Consequently, the market mechanism would direct the price lower (toward equilibrium) until it falls to P_E. In this case, equilibrium is **dynamically stable**. If market price were lower than equilibrium price, there would be a shortfall so the market mechanism would drag the price higher (toward equilibrium) until it climbs to P_E.

On the other hand, in Figure 2-2b, it is the demand curve that is steeper, and the supply curve intersects the demand curve from below. At prices greater than equilibrium price (P_E), for example at P_H, quantity supplied (QS_G) is less than quantity demanded (QD_G) so there is a shortfall. Consequently, the market mechanism would take the price higher dragging it further away from equilibrium price. Similarly, at a price lower than P_E, there would be excess supply and again the market mechanism would take the price further away from P_E by inducing it to fall. Such equilibrium is **dynamically unstable.**

It is also possible for markets to have multiple equilibria, as illustrated in Figure 2-3. The demand and supply curves intersect at $P_{E\text{-high}}$ and then again at $P_{E\text{-low}}$. $P_{E\text{-high}}$ is a dynamically unstable equilibrium as both the demand and supply curves are negatively sloped at this point and supply intersects demand from below (similar to Figure 2-2b). On the other hand, $P_{E\text{-low}}$ is a dynamically stable equilibrium as supply is upward-sloping and demand is downward sloping at this point. Interestingly, note that if prices fall below $P_{E\text{-high}}$, there would be excess supply, which would push prices even lower (away from $P_{E\text{-high}}$ but toward the stable equilibrium, $P_{E\text{-low}}$).

Figure 2-3: Multiple Equilibria



Non-linear supply curves can result in multiple equilibria.

LOS 13i: Describe types of auctions and calculate the winning price(s) of an auction.

Auctions as a Way to Find Equilibrium Price

Auctions can be classified into two types based on whether or not the value of the item to each bidder is the same.

Common value auction: The value of the product is the **same** to each bidder. Bidders estimate the value of the product before the auction is settled and the common value of the product is revealed once the auction is complete. For example, consider an auction for a jar full of coins where bidders must estimate the value to submit bids in the auction, and the actual value is only revealed once the auction has been settled.

Private value auction: Each bidder places a subjective value on the product, and the value each bidder places on the product is generally **different**. For example, consider an auction for a unique piece of art where buyers place bids depending on how much they value the item.

Auctions also differ with respect to how the **final price** and **eventual buyer** are determined.

Ascending price auction: Potential buyers openly reveal their bids at prices that are called out by the auctioneer. The auctioneer starts the bidding at a particular price and then raises the price in response to nods from bidders. In a common value auction, buyers may be able to learn more about the true value of the item from the bids placed by other potential buyers. As the price rises, bidders begin to drop out until only one bidder is left. The item is sold to her for the last price that she bid.

Sealed bid auction: In a sealed bid auction for a common value item (e.g., timber lease) potential buyers bid for the item with no knowledge of the values bid by other potential buyers (until after the auction has been settled).

In a first price sealed bid auction, all envelopes containing bids are opened simultaneously and the item is sold to the highest bidder for the price she bid. Further, there is no certainty regarding expected profits at the time of bidding; eventual profits are only known once the asset is exploited. This opens up the possibility of the winner's curse where a buyer purchases the asset for a price greater than the value eventually realized from the asset. Therefore, bidders tend to be overly cautious with their bids in such auctions, which might result in the seller attaining a relatively low price for the asset.

If the item being sold is a private value item, there is no danger of the winner's curse as no one would bid more than her own valuation. However, bidders will try to guess the reservation prices of other bidders, as the optimal outcome for the successful bidder would be to win the auction with a bid just above the reservation price of the second highest bidder. In order to induce bidders to reveal their true reservation prices, sellers can use a second price sealed bid auction (also known as a Vickery auction). Sealed bids are opened simultaneously, the highest bidder wins the bid, but the price paid by the winner equals the second-highest bid (not her own bid). The optimal strategy for each bidder in such an auction is to bid her true reservation price, so the second price sealed bid auction induces bidders to reveal their true valuation of the item. Note that a second price sealed bid auction will result in the same outcome as an ascending price auction (as long as bidding increments are small).

The reservation price represents the highest amount that a bidder is willing to pay for the item.

Descending price or Dutch auction: The bidding begins at a very high price and the auctioneer lowers the price in increments until the item is sold. Suppose that only one unit of the item is being offered for sale. The supply curve for the item will be perfectly inelastic at a quantity equal to 1 unit and would extend from the seller's reservation price upward. Each bidder's demand curve is perfectly inelastic at a quantity equal to 1 unit and will extend up to her reservation price. The market demand curve will be a negatively sloped step function. The highest/left-most point on the demand curve would have a coordinate of 1 unit and the highest bidder's reservation price. An increase in quantity demanded of 1 unit would require a reduction in price to the next-highest bidder's reservation price, and so on.

A Dutch auction can have a single-unit (as described above) or a multiple-unit format. In a multiple-unit format, the price quoted by the auctioneer is a per-unit price, and the winning bidder can purchase as many units as she wants at that price. If the winning bidder purchases less than all the units available for sale, the auctioneer drops the price in increments until eventually all units are sold. Therefore, multiple transactions can occur at multiple prices.

A modified Dutch auction (widely used in securities markets) establishes one price for all purchases. Stock repurchases are conducted using this method, where the aim is to establish the minimum price at which the company can repurchase all the shares it wants to repurchase. Each successful bidder is then paid this price. For example, a company may offer to repurchase 10 million shares in a range of $50 to $55 per share. Each shareholder would submit the number of shares she wants to put up for sale and specify the price she is willing to accept for her shares. The company would begin qualifying bids starting with those offered at $50 and proceed upward until 10 million shares have qualified. If 10 million shares are qualified at a price of $52.75, then all shareholders who bid between $50 and $52.75 (both prices inclusive) would have their specified quantities of shares repurchased by the company at the rate of $52.75 per share.

Another variation of the Dutch auction, known as a single price auction, is used in the U.S. Treasury market. The Treasury puts up Treasury bills for sale with both competitive and non-competitive bidding. Non-competitive bidders simply state the total face value they are willing to purchase at the final price (yield) that clears the market (results in all the securities on offer being sold). Competitive bidders specify the total par value they want to purchase and the exact price (yield) at which they are willing to purchase that quantity. The Treasury then ranks the bids in descending order of price (ascending order of yield) and determines the price (yield) at which the market would clear. Example 2-1 illustrates how a single price auction works.

Example 2-1: Single Price Auction

The U.S. Treasury offers to sell $120 billion worth of 26-week T-Bills and requests competitive and non-competitive bids. It receives non-competitive bids worth $20 billion. Competitive bids received are listed in descending order of price:

Discount Rate Bid (%)	Bid Price per $100	Competitive Bids ($ billion)
0.1685	99.8345	15
0.1690	99.8340	12
0.1695	99.8335	3
0.1700	99.8330	25
0.1705	99.8325	35
0.1710	99.8320	25
0.1715	99.8315	10

1. Determine the winning price if a single price auction is used to sell the T-Bills.
2. For those bidders who bid exactly the winning price, what percentage of their orders would actually be filled?

Solution:

The first step in determining the winning price of a single price auction is to organize the bids in ascending order of yield (descending order of price). The table in the question has already done that for us. Next, we determine cumulative competitive bids and then cumulative total bids (add non-competitive bids to cumulative competitive bids). These calculations have been performed in the table:

Discount Rate Bid (%)	Bid Price per $100	Competitive Bids ($ billions)	Cumulative Competitive Bids ($ billions)	Non-Competitive Bids ($ billions)	Total Cumulative Bids ($ billions)
0.1685	99.8345	15	15	20	35
0.1690	99.8340	12	27	20	47
0.1695	99.8335	3	30	20	50
0.1700	99.8330	25	55	20	75
0.1705	99.8325	35	90	20	110
0.1710	99.8320	25	115	20	135
0.1715	99.8315	10	125	20	145

Notice that at yields lower than 0.1710%, there is excess supply. For example, at a yield of 0.1705, total cumulative bids amount to $110 billion, while the quantity offered for sale equals $120 billion. At this yield (0.1710%) however, total cumulative bids ($135 billion) exceed the total size of the offer ($120 billion). Therefore, 0.1710% is known as the clearing yield and all sales will be made at this yield. First, all non-competitive bidders will have their orders (worth total par value of $20 billion) filled at this yield. All competitive bidders who bid a lower yield (higher price) will also have their orders (worth total par value of $90 billion) filled at this yield. The remaining $10 billion of the offer (total size of offer ($120 billion) – Non-competitive bids ($20 billion) – Cumulative competitive bids at yields lower than the clearing yield ($90 billion) are distributed pro rata among buyers who bid the clearing yield. Therefore, buyers who bid the clearing yield would have 10/25 × 100 = 40% of their orders filled.

LESSON 3: CONSUMER SURPLUS AND PRODUCER SURPLUS: INTRODUCTION AND APPLICATIONS

LOS 13j: Calculate and interpret consumer surplus, producer surplus, and total surplus.

The concepts of consumer and producer surplus allow us to evaluate whether the outcome of a competitive market is actually socially optimal.

The Demand Curve, Value (Utility), and Consumer Surplus

The utility derived from consumption of the last unit of a good or service is known as its **marginal benefit (MB)**. Consumers express how much utility they derive from consuming an additional unit of a good or service through the price they are willing to pay for it. If a bottle of soda gets Alan $5 worth of utility, he will only be willing to pay $5 for it. In other words, the marginal benefit curve is the demand curve where the benefit or utility that we derive from the good is quantified by the maximum price we are willing and able to pay for it. Marginal benefit is downward sloping because the utility derived from consumption of the next unit will be lower than the utility derived from consumption of the last unit (law of diminishing marginal utility), and therefore a consumer would be willing to pay a lower price for each additional unit of a given product.

The CFA Program curriculum calls the demand curve the marginal value curve. We use the term marginal benefit curve to describe the demand curve.

We are measuring utility here in terms of the price that a consumer is willing and able to pay for each soda bottle (in dollars). Alternatively, the price reflects the value of other goods and services that a consumer is willing to forego the consumption of, in order to consume one more bottle of soda.

Consumer Surplus

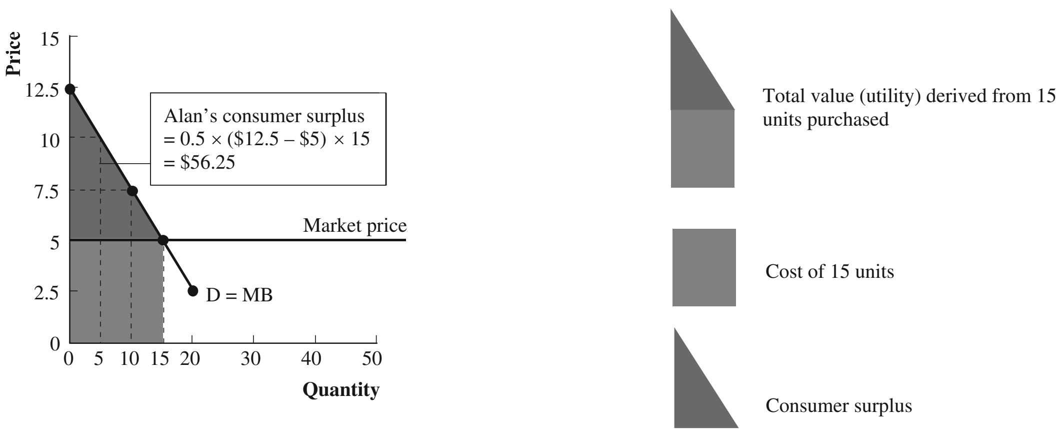
Consumer surplus occurs when a consumer is able to purchase a good or service for less than the maximum price that she is willing and able to pay for it. It equals the difference between the price that a consumer is willing and able to pay for a good (indicated by her demand curve) and what she actually pays for the good (the market price, which is determined through the interaction of demand and supply in the market). The existence of a free market allows consumers to pay a uniform market price for each additional unit purchased regardless of the value they place on consumption of that unit of the good.

Figure 3-1 shows Alan's consumer surplus from consuming bottles of soda at a price of $5/bottle. At the market price of $5, he purchases 15 bottles and his marginal benefit from consuming the 15^{TH} unit is the same as the market price. Notice that Alan is willing and able to pay $10 for the 5^{TH} bottle, but only pays $5 (the market price) for it. This implies that his marginal benefit from consuming the 5^{TH} unit exceeds its price by $5. Alan essentially receives a "bargain" on his purchase as he pays $5 less than what he was willing and able to pay. This $5 represents Alan's consumer surplus on the 5^{TH} unit. Similarly, the value Alan places on consumption of the 10^{TH} unit is the price he is willing and able to pay for it ($7.50). He is also able to purchase the 10^{TH} unit at the market price ($5) so his consumer surplus on consumption of the 10^{TH} unit equals $7.50 – $5 = $2.50. Alan's total consumer surplus is the sum of all the surpluses that he incurs from the purchase of each bottle of soda that he consumes, up to and including the last unit (15^{TH} unit). Therefore, Alan's total consumer surplus is the area of the triangle among his demand curve, the vertical axis, and the market price.

Note that the total value or utility that Alan derives from consuming the 15 units of soda (in terms of dollars) equals the trapezoid (green triangle plus grey rectangle) under the demand curve between the vertical axis and total units consumed (15). Alan incurs a total cost of consuming these 15 units (at $5 each) equal to the area of the rectangle shaded in grey. The difference between the total value derived from consumption and the total cost of purchase equals consumer surplus.

Formally, we can define consumer surplus as the difference between the value consumers place on units purchased and the amount of money that was required to purchase them.

Figure 3-1: Demand and Consumer Surplus



Alan's Consumer Surplus

The Supply Curve, Marginal Cost, and Producer Surplus

> We examine the relation between the supply curve and the marginal cost curve in detail in a later reading.

The cost of producing one more unit of output is known as **marginal cost (MC)**. A producer will not be willing to supply an additional unit of a good if the price she expects to receive from the sale is lower than the marginal cost of producing it. A producer will only be willing to supply an additional unit of a product when the price she expects to receive for the unit exceeds its marginal cost as the excess of price over marginal cost would serve to meet fixed costs and contribute to profits. An individual firm's supply curve illustrates its willingness and ability to produce a good at various prices. Therefore, an individual firm's supply curve is essentially its marginal cost curve.

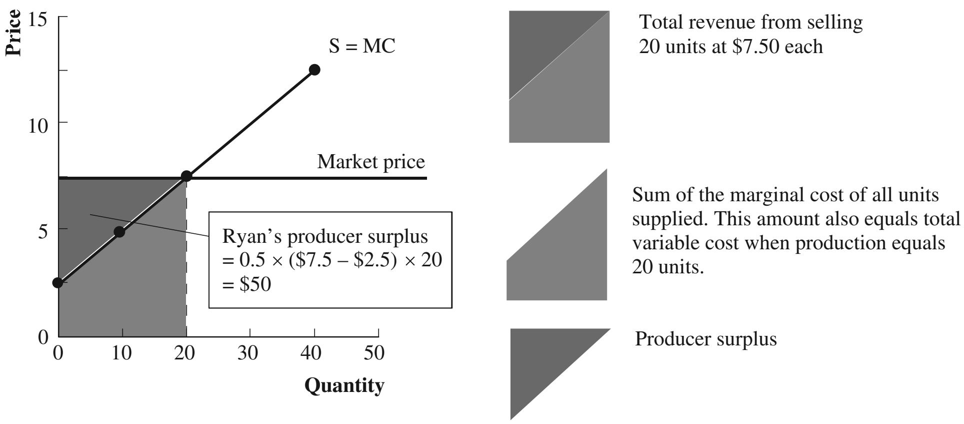
Marginal cost curves are typically upward-sloping due to the law of diminishing marginal returns (which we will also explain in a later reading).

Producer Surplus

Producer surplus occurs when a supplier is able to sell a good or service for more than the price that she is willing and able to sell it for. It equals the difference between the market price and the price at which producers are willing and able to sell their product (marginal cost), which is indicated by the supply curve. In Figure 3-2, at a market price of $7.50, Ryan produces 20 bottles because the marginal cost of producing the 20TH bottle equals market price ($7.50). Notice that Ryan is willing to sell the 10TH bottle for only $5 (as the marginal cost of producing it equals $5), but due to the existence of a standard market price, he will actually get $7.50 for it. This excess of price over the marginal cost of the 10TH unit ($7.50 – $5 = $2.50) is Ryan's producer surplus from supplying the 10TH unit. Ryan's total producer surplus is the sum of the individual surpluses that he earns from selling each bottle, up to and including the last unit supplied (the 20TH bottle). Therefore, his total producer surplus is the area of the green triangle between his supply curve, the vertical axis and market price.

Note that the total revenue from selling 20 units (at $7.50 each) equals the area of the rectangle (regions shaded in green and grey combined) in Figure 3-2. The total variable cost (sum of the marginal costs of producing each of the 20 units) equals the area of the region shaded in grey. The difference between total revenue and total variable cost equals producer surplus (area shaded in green).

Figure 3-2: Supply and Producer Surplus

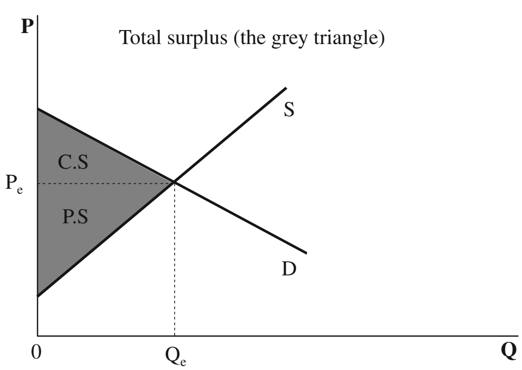


Ryan's Producer Surplus

Total Surplus: Total Value (Utility) Minus Total Variable Cost

We have now seen that the existence of an equilibrium market price benefits both buyers and sellers. For every unit up to and including the equilibrium unit, buyers are able to purchase for less than they were willing and able to pay, while sellers are able to sell for more than they were willing and able to accept. The total value to buyers was greater than the total variable cost to sellers. The difference between these two quantities is called total surplus. Another way to look at total surplus is the sum of producer and consumer surplus (see Figure 3-3).

Figure 3-3: Total Surplus



The distribution of total surplus between consumers and producers depends on the relative slopes of the demand and supply curves. If the supply curve is steeper, more of the surplus is captured by producers. If the demand curve is steeper, more of the surplus is captured by consumers.

Total surplus can be looked upon as society's gain from the existence of a free market where goods can be exchanged voluntarily. An important thing for us to know is that total surplus is maximized at equilibrium where the market supply and demand curves intersect.

An externality is when a production cost or consumption benefit spills over to those not producing or consuming the service. Pollution is an example of a negative externality, while education has positive externalities.

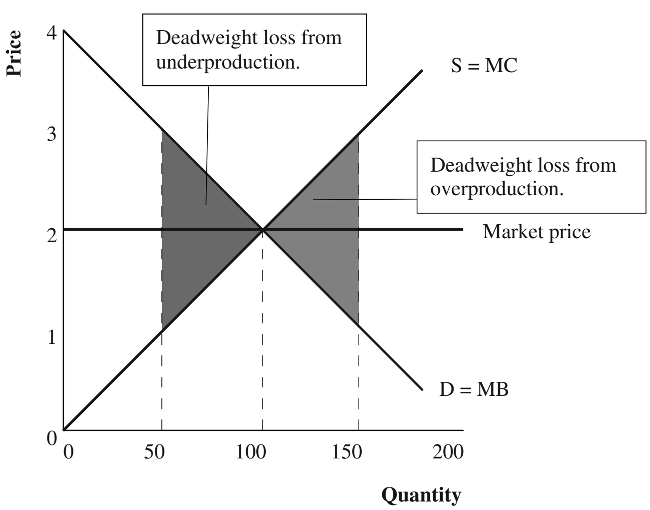
Markets Maximize Society's Total Surplus

The market demand curve represents a society's marginal value or marginal benefit curve for a particular good, while the market supply curve represents the marginal cost to society of producing each additional unit of the good, assuming no externalities. At equilibrium (where the market demand and supply curves intersect) the highest price someone is willing and able to pay for the good (MB) equals the lowest price a seller is willing to accept (MC).

In Figure 3-4, equilibrium occurs at a quantity of 100 bottles of soda. Let's examine a situation where only 50 bottles of soda are being traded. At this point consumers are willing to pay \$3 for the 50^{TH} bottle of soda, which is produced at a cost of \$1. The marginal benefit of consuming the 50^{TH} unit is greater than the marginal cost of producing it. Society should produce the 50^{TH} unit, as well as the next, and the next all the way up to 100 units. Limiting total output to 50 units results in a reduction in the sum of consumer and producer surplus (total surplus is not maximized). This combined reduction in consumer and producer surplus is known as a **deadweight loss**, which is borne by society as a whole. The deadweight loss from underproduction is the region shaded in green in Figure 3-4.

Figure 3-4 also illustrates a situation where 150 bottles of soda are being supplied. The 150^{TH} bottle of soda costs \$3 to produce, but consumers are only willing to pay \$1 for it. The marginal bottle of soda costs more than the value consumers place on it. This results in inefficiency as too many soda bottles are being produced (i.e., the marginal cost of producing soda bottles is too high). Society should not have produced and consumed all the additional units beyond 100 units. The deadweight loss from overproduction is the area shaded in grey in Figure 3-4.

Figure 3-4: Deadweight Losses

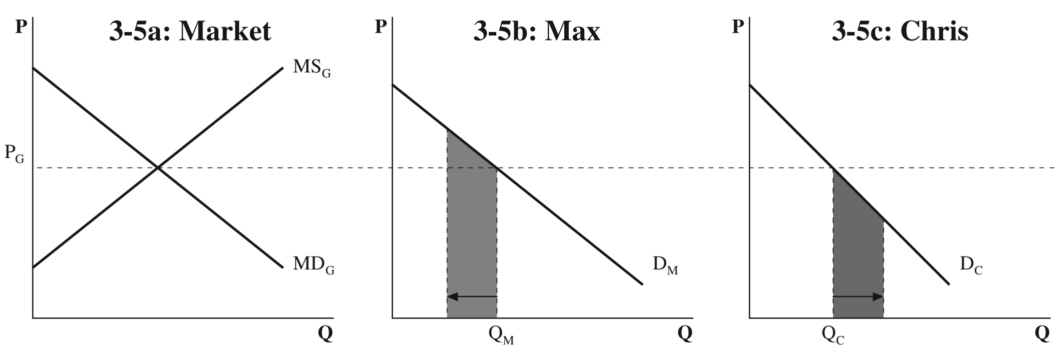


We have already seen that the market mechanism tends to pull the market toward equilibrium. We have now learned that the condition of equilibrium is optimal in terms of maximizing total welfare. Let's examine this second statement more closely by looking at individual consumers. Figure 3-5a presents the market demand curve (MD_G) and supply curve (MS_G) for gasoline. The point of intersection of these two curves determines the price of gasoline (P_G). Figures 3-5b and 3-5c present the individual demand curves for Max and Chris (two individual consumers). At the market price of P_G, Max purchases Q_M and Chris purchases Q_C units because at this price the marginal value derived from consumption of the last unit for each of them equals the price that they have to pay for that unit.

Now suppose that one unit (from Q_M) was taken away from Max and given to Chris (in addition to Q_C). The grey trapezoid depicts the loss in value (utility) incurred by Max from consuming one less unit, while the green trapezoid depicts the gain in value (utility) that accrues to Chris from consuming the additional unit. Notice that the loss in value experienced by Max (area in grey) is greater than the gain in value experienced by Chris (area in green). Total value is reduced when individuals consume quantities that do not yield equal marginal value (utility) to each of them. Expenditure remains the same as the same total number of units is being purchased by society. Therefore, consumer surplus falls (recall that consumer surplus equals value minus expenditure). When all consumers face the same price, consumer surplus is maximized when each consumer purchases a quantity that equates her marginal value from the last unit consumed to the market price.

Similarly, producer surplus is maximized when each producer supplies a quantity that equates the marginal cost from the last unit supplied to the market price. To conclude, we state that when there is a single price that is freely determined in the market, consumers have the opportunity to purchase all they want at this price, and producers have the opportunity to sell all they want at this price, and total surplus is maximized.

Figure 3-5



LOS 13k: Describe how government regulation and intervention affect demand and supply.

LOS 13l: Forecast the effect of the introduction and removal of a market interference (e.g., a price floor or ceiling) on price and quantity.

Market Intervention: Negative Impacts on Total Surplus

Price Ceilings

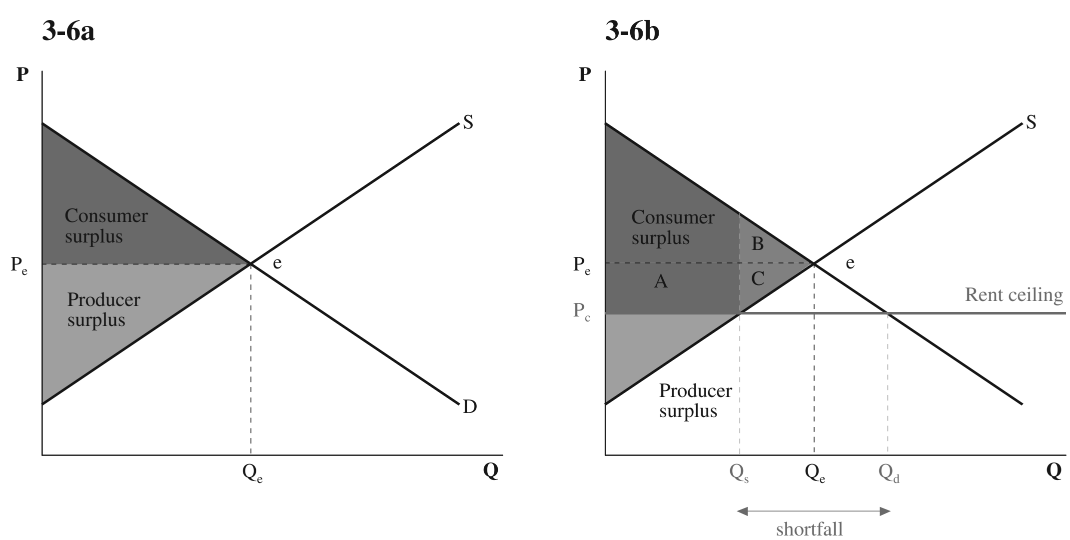
There are times when the government may feel that the market price for a good or a service is too high, so it may be tempted to impose a ceiling or limit on the price below the equilibrium market price. An example of such a ceiling is rent control. On the face of it, limiting rental rates to more affordable rates (than currently prevailing in the market) seems like a noble idea, but the truth is a bit more complicated.

Suppose the housing market is currently in equilibrium at Point e (see Figure 3-6a), with quantity = Q_e and price = P_e. Total consumer surplus equals the area of the triangle shaded in dark green, while total producer surplus equals the area of the triangle shaded in light green. Total surplus equals the sum of these two triangles and since we are operating at free market equilibrium, total surplus is maximized at Point e.

The government then decides to impose a rent ceiling below the market price at a level of P_c (see Figure 3-6b). At this lower price, quantity demanded (Q_d) exceeds quantity supplied (Q_s) so there is a shortfall. Only a quantity of Q_s is traded in the market, and this quantity is less than the quantity that was being traded before the imposition of the ceiling (Q_e). Consumers gain in the sense that those who are able to purchase some of the Q_s units for sale obtain those units at a lower price (P_c versus P_e earlier) so some of the producer surplus (Rectangle A) gets transferred to consumers. This region continues to be a part of total surplus.

Price ceilings only disrupt market equilibrium if set below the equilibrium market price. They have no effect on economic activity if set above the equilibrium market price.

Figure 3-6: A Price Ceiling



However, there is a loss in consumer surplus (Triangle B) and producer surplus (Triangle C) because there is no action in the market beyond Q_s units, which is less than the previous quantity (Q_e). The loss of these surpluses is known as a deadweight loss (surplus lost by one group that is not transferred to another). Consumer and producer surplus after imposition of the ceiling are shaded in dark green and light green respectively on Figure 3-6b.

Price Floors

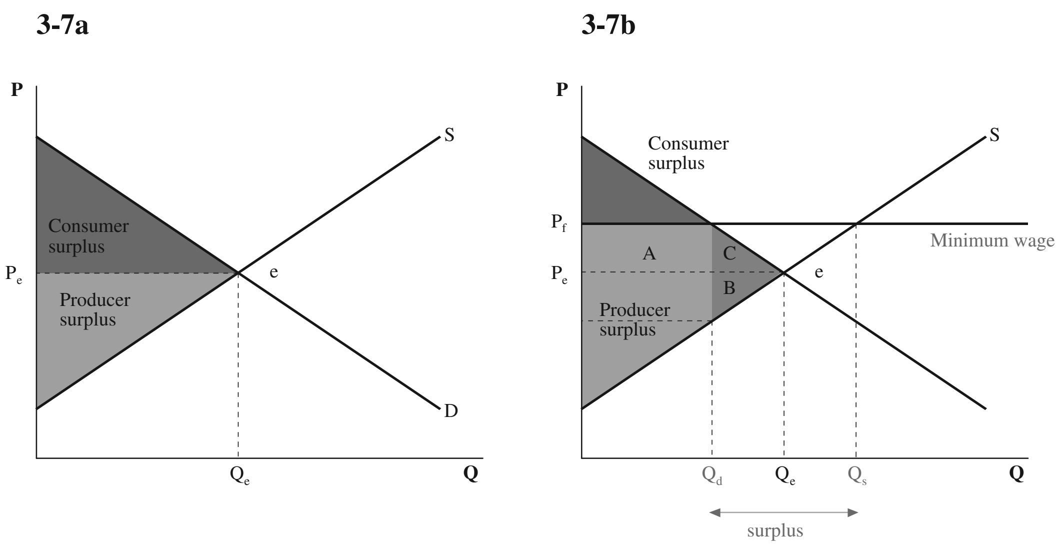
At other times, a government may feel that the market price for a good or a service is too low, so it may be tempted to impose a minimum price above the equilibrium market price. An example of such a floor is the minimum wage. On the face of it, setting wages at higher rates (than currently prevailing in the market) seems like a noble idea, but again the truth is a bit more complicated.

Suppose the labor market is currently in equilibrium at Point e (see Figure 3-7a), with quantity = Q_e and price = P_e. Total consumer surplus equals the area of the triangle shaded in dark green, while total producer surplus equals the area of the triangle shaded in light green. Total surplus equals the sum of these two triangles and since we are operating at free market equilibrium, total surplus is maximized at Point e.

> Price floors only disrupt market equilibrium if set above the equilibrium market price. They have no effect on economic activity if set below the equilibrium market price.

The government then decides to impose a minimum wage above the market price at a level of P_f (see Figure 3-7b). At this higher wage, quantity supplied (Q_s) exceeds quantity demanded (Q_d) so there is a surplus. Only a quantity of Q_d is traded in the market, and this quantity is less than the quantity that was being traded before the imposition of the floor (Q_e). Suppliers gain in the sense that those who are able to sell some of the Q_d units purchased sell those units at a higher price (P_f versus P_e earlier) so some of the consumer surplus (Rectangle A) gets transferred to suppliers. This region continues to be a part of total surplus. However, there is a loss in producer surplus (Triangle B) and consumer surplus (Triangle C) because there is no action in the market beyond Q_d units, which is less than the welfare-maximizing quantity. These two triangles comprise the deadweight loss from imposition of the minimum wage. Consumer and producer surplus after imposition of the floor are shaded in dark green and light green respectively on Figure 3-7b.

Figure 3-7: A Price Floor



Taxes

Per-Unit Tax on Sellers

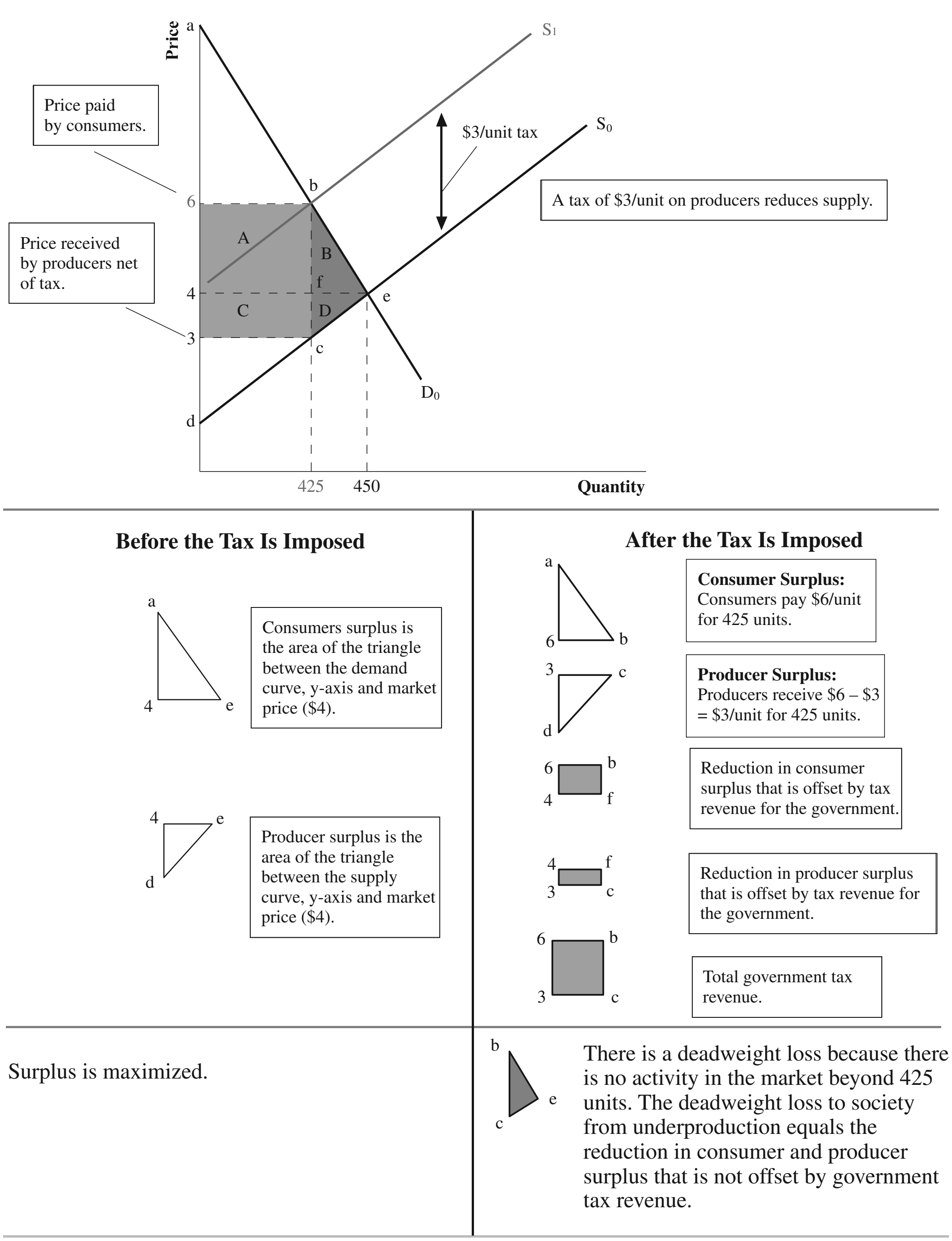
The statutory incidence of a tax refers to whom the law levies the tax upon. As we shall learn soon, just because the government imposes or levies a tax on a particular group does not necessarily mean that the actual incidence of the tax falls entirely on that group. Actual tax incidence refers to how the burden of the tax is shared by consumers and producers in terms of reductions in consumer and producer surplus respectively.

Let's start with an example in which a tax per unit of $3 is levied on suppliers (see Figure 3-8). The supply curve presents the minimum prices that producers are willing and able to sell various quantities of output for. Since they must now pay $3 in taxes on each unit sold, suppliers will sell every given quantity for $3 higher (so that they still cover their marginal cost on each unit sold). As a result, supply shifts to the left to S_1. Consequently, equilibrium price rises to $6 and equilibrium quantity falls to 425 units.

- Consumers purchase 425 units and pay $6/unit. Effectively prices paid by consumers have gone up by $6 – $4 = $2. Consumer surplus has therefore fallen by Rectangle A and Triangle B.
- Producers sell 425 units at $6/unit but only pocket $3/unit after paying the tax. Effectively, their realized prices have fallen by $4 – $3 = $1. Producer surplus has therefore fallen by Rectangle C and Triangle D.
- The government earns tax revenue of $3/unit on 425 units that are sold. So part of the loss in consumer surplus (Rectangle A) and producer surplus (Rectangle C) is transferred to the government.
- However, some consumer surplus (Triangle B) and producer surplus (Triangle D) remains untransferred and is lost due to the imposition of the tax. These two triangles comprise society's deadweight loss.

Even though this tax was levied on suppliers only, consumers and producers share the actual burden of the tax as consumer and producer surplus both decline once the tax is imposed. Further, in our example, consumers actually end up bearing the brunt of the tax in the form of an effective increase in prices of $2, versus an effective decrease in producer realized prices of only $1. Note that consumer surplus transferred to the government, Rectangle A, is greater than producer surplus transferred to the government, Rectangle C. This is because the demand curve is steeper than the supply curve. If the supply curve were steeper, the reverse would be true regardless of whom the tax was imposed upon by law.

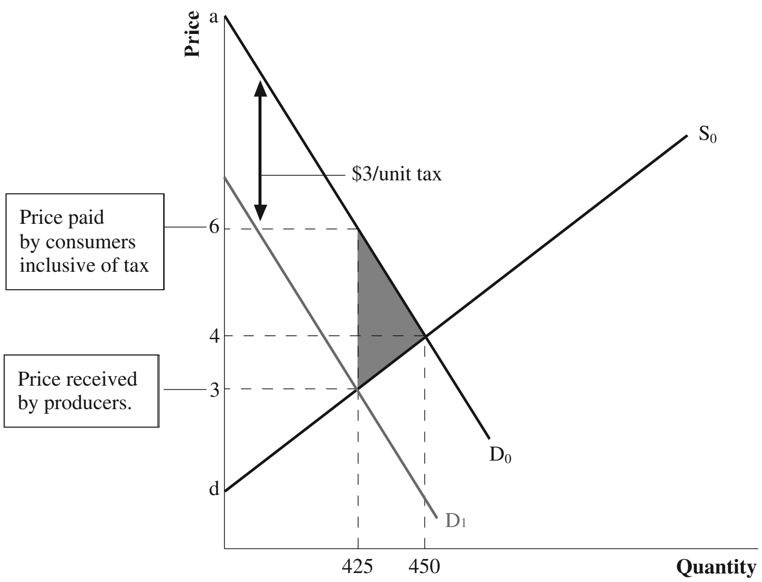
Figure 3-8: Tax on Sellers



Tax on Buyers

Now assume that instead of being levied upon producers, the same $3/unit tax is imposed on consumers (see Figure 3-9). The demand curve shifts to the left to D_1. The actual burden of the tax is shared by consumers and producers, and once again (since the demand curve is steeper than the supply curve) consumers bear a greater burden of the tax. Further, the increase in government revenue from tax collections does not entirely offset the reduction in consumer and producer surplus, and society suffers a deadweight loss from underproduction (the region shaded in grey).

Figure 3-9: Tax on Consumers



Aside from price ceilings, price floors, and taxes, other examples of governments interfering with the free market mechanism include tariffs, quotas, and bans on imports. Governments may also regulate consumption or production of goods that have negative effects on third parties (externalities). For example, the government may limit the production of an item that significantly pollutes the environment. Government interference in such markets (where market prices do not account for externalities) seems justified. However, in markets where social marginal benefit and social marginal costs are truly reflected in the market demand and supply curves respectively (i.e., there are no externalities) total surplus is maximized if markets are allowed to operate freely.

Search Costs

The costs of matching buyers and sellers in the market are known as search costs. There may be a buyer who is willing and able to purchase a product for a price greater than the minimum price that a producer is willing and able to accept, but if the two do not find each other, the transaction will not be completed, societal surplus will not be maximized, and a deadweight loss will result.

When search costs are significant, brokers can play a valuable role in the market by bringing buyers and sellers together. Brokerage costs may be looked upon negatively as transaction costs, but a more appropriate way of viewing them would be to consider them as the price of reducing search costs. The service provided by brokers improves information flow and facilitates the exchange of goods and services, with every exchange contributing to total surplus. Similarly, informative advertising can also add value to the extent that it informs market participants about the availability of goods and services.

LESSON 4: DEMAND ELASTICITIES

LOS 13m: Calculate and interpret price, income, and cross-price elasticities of demand and describe factors that affect each measure.

Demand Elasticities

Own-Price Elasticity of Demand

A firm's total revenue equals quantity sold times price. Given that price and quantity demanded are negatively related, a firm needs to know how sensitive quantity demanded is to changes in price to determine the overall impact of a price change on total revenue. For example, if prices were increased, a firm would need to analyze how much quantity demanded would fall and how total revenue would be affected. If the percentage increase in price is greater than the percentage decrease in quantity demanded, total revenue will increase. If the percentage increase in price is lower than the percentage decrease in quantity demanded, total revenue will decline.

Let's go back to Equation 12 (market demand function):

$$QD_G = 12{,}000 - 500P_G \quad \text{... (Equation 12)}$$

One measure of the sensitivity of quantity demanded to changes in price is the slope of the demand function. Equation 12 tells us that for a one-unit change in price, quantity demanded moves by 500 units in the other direction. Unfortunately, this measure is dependent on the units in which we measure QD and P. Therefore, economists prefer to use elasticity as a measure of sensitivity. Elasticity uses percentage changes in the variables and is independent of the units used to measure the variables.

The own-price elasticity of demand is calculated as:

$$ED_{Px} = \frac{\%\Delta QD_x}{\%\Delta P_x} \quad \text{... (Equation 16)}$$

If we express the percentage change in X as the change in X divided by the value of X, Equation 16 can be expanded to the following form:

$$ED_{Px} = \frac{\%\Delta QD_x}{\%\Delta P_x} = \frac{\Delta QD_x / QD_x}{\Delta P_x / P_x} = \left(\frac{\Delta QD_x}{\Delta P_x}\right)\left(\frac{P_x}{QD_x}\right) \quad \text{... (Equation 17)}$$

Slope of demand function

Coefficient on own-price in market demand function

The expression above tells us that we can compute the own-price elasticity of demand by multiplying the slope (–500) of the demand function, $\Delta QD_x/\Delta P_x$ (or the inverse of the slope of the demand curve, $1/[\Delta QD_x/\Delta P_x]$), by the ratio of price to quantity, P_x/QD_x. At a price of \$2/gallon, using the equation for the market demand for gasoline ($QD_G = 12,000 - 500\ P_G = 11,000$), we can compute the own-price elasticity of demand for gasoline as:

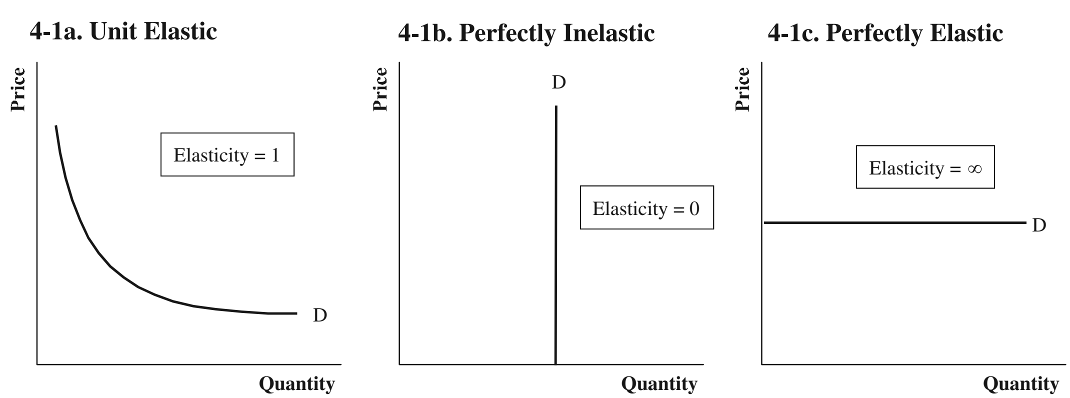
Recall that the slope of the demand curve equals the slope of the inverse demand function.

$$ED_{PG} = -500 \times (2/11,000) = -0.09091$$

We usually look at the absolute value of own-price elasticity of demand when classifying how sensitive quantity demanded is to changes in price:

- If own-price elasticity of demand equals 1 (percentage change in quantity demanded is the same as the percentage change in price) demand is said to be unit elastic (see Figure 4-1a).
- If own-price elasticity of demand equals 0 (quantity demanded does not change at all in response to a change in price) demand is said to be perfectly inelastic (see Figure 4-1b).
- If own-price elasticity of demand equals ∞ (quantity demanded changes by an infinitely large percentage in response to even the slightest change in price) demand is said to be perfectly elastic (see Figure 4-1c).
- If the absolute value of price elasticity of demand lies between 0 and 1, demand is said to be relatively inelastic.
- If the absolute value of price elasticity of demand is greater than 1, demand is said to be relatively elastic.

Figure 4-1: Price Elasticity of Demand

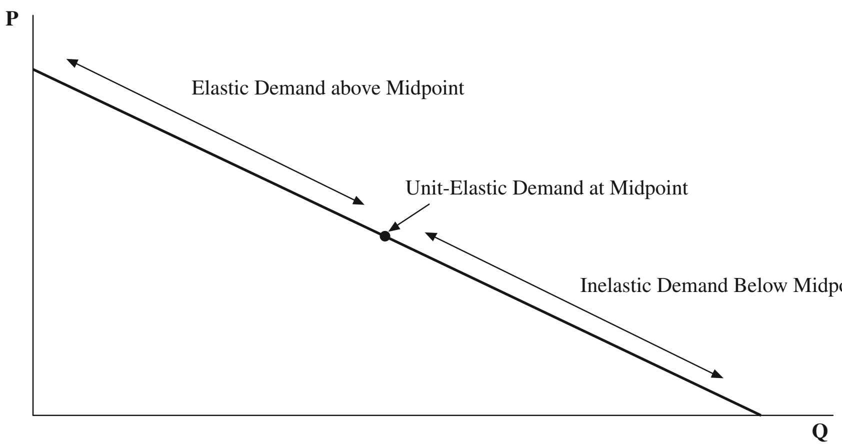


In our example, demand for gasoline appears to be relatively inelastic as the absolute value of own-price elasticity lies between 0 and 1.

An important thing to note is that while the slope of the demand curve remains constant along a downward-sloping linear demand curve, the ratio of price to quantity is different at each point along the demand curve (unless the demand curve is perfectly elastic or perfectly inelastic), as shown in Figure 4-2.

- At relatively low prices (relatively high quantities) the ratio of price to quantity is relatively low so own-price elasticity of demand (absolute value of ED_P) is low and demand is relatively inelastic.
- At relatively high prices (relatively low quantities) the ratio of price to quantity is relatively high so own-price elasticity of demand (absolute value of ED_P) is high and demand is relatively elastic.
- Demand is unit elastic at the midpoint of the demand curve, relatively elastic above the midpoint, and relatively inelastic below the midpoint.

Figure 4-2: The Elasticity of a Linear Demand Curve



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

When information on the entire demand curve is not available, but any two observations of price and quantity demanded are available, arc elasticity may be used to gauge the responsiveness of quantity demanded to changes in price. Arc elasticity is calculated as:

$$E_P = \frac{\text{\% change in quantity demanded}}{\text{\% change in price}} = \frac{\%\Delta\ Q_d}{\%\Delta\ P} = \frac{\dfrac{(Q_0 - Q_1)}{(Q_0 + Q_1)/2} \times 100}{\dfrac{(P_0 - P_1)}{(P_0 + P_1)/2} \times 100} \quad \text{... (Equation 18)}$$

Note that changes in quantity demanded and price are calculated using *average* amounts in the denominator. This method results in the same value for elasticity whether prices move up or move down to a particular level.

Example 4-1: Price Elasticity of Demand

If the price of a product decreases from \$7 to \$6, its quantity demanded increases from 18 to 20 units. What is the price elasticity of demand for this product?

Solution

$$E_P = \frac{\frac{(20-18)}{(20+18)/2} \times 100}{\frac{(6-7)}{(6+7)/2} \times 100} = \frac{(2/19) \times 100}{(-1/6.5) \times 100} = 0.684$$

Factors Affecting Own-Price Elasticity of Demand

Availability of Close Substitutes

If a consumer can *easily* switch away from a good, her ability to respond to a price increase (by reducing consumption of the good) is *high*, and demand for that product would be *relatively elastic*. Generally speaking, the demand curve faced by an individual producer is relatively more elastic than the demand curve for the entire market. For example, demand for Nike® shoes is more elastic than demand faced by the shoe industry, as there are more substitutes for Nike® shoes (e.g., Reebok®, Adidas®, etc.) than for shoes in general.

Proportion of Income Spent on the Good

If a relatively small proportion of a consumer's income is spent on a good (e.g., soap), she will not significantly cut down on consumption if prices increase. Demand for such a good will be relatively inelastic. However, if consumption of the good takes up a larger proportion of her income (e.g., automobiles), she might be forced to reduce quantity demanded significantly when the price of the good increases. Demand for such a good will be relatively elastic.

Time Elapsed Since Price Change

The *longer* the time that has elapsed since the price change, the *more elastic* demand will be. For example, if the price of labor goes up, firms may not be able to make radical changes to their production methods in the short term and therefore, demand for labor will be *relatively inelastic*. However, in the long term, if the price of labor remains high, firms may automate their production processes and substitute machinery for labor. In the long run, demand for labor will be *relatively elastic*.

The Extent to Which the Good is Viewed as Necessary or Optional

The more the good is seen as being necessary, the less elastic its demand is likely to be. For example, demand for milk is less elastic than demand for opera tickets.

Own-Price Elasticity of Demand and Total Expenditure

We established earlier that own-price elasticity changes along the demand curve. Now let's look into how total expenditure on a good changes as its price fluctuates.

Total expenditure (and revenue) equals price times quantity purchased (sold). If prices are *reduced* to stimulate sales, total revenue will only increase if the percentage *increase* in demand (sales) is *greater* than the percentage *decrease* in prices.

The relationship between total expenditure and price depends on price elasticity of demand:

- If demand is relatively elastic (elasticity greater than 1), a 5% *decrease* in price will result in an *increase* in quantity demanded of *more* than 5%. Therefore, total expenditure will *increase*.
- If demand is relatively inelastic (elasticity less than 1), a 5% *decrease* in price will result in an *increase* in quantity demanded of *less* than 5%. Therefore, total expenditure will *decrease*.
- If demand is unit elastic, a 5% *decrease* in price will result in an *increase* in quantity demanded of exactly 5%. Therefore, total expenditure will *not change*.

When price elasticity is greater than one (relatively elastic), the numerator MUST be greater than the denominator. If the denominator changes by 5%, the numerator HAS to change by more than 5%. The effect of a decrease in prices will be outweighed by the effect of the increase in quantity demanded and total expenditure will rise.

The total expenditure (revenue) test gauges price elasticity by looking at the direction of the change in total revenue in response to a change in price:

- If the price cut *increases* total revenue, demand is relatively elastic.
- If the price cut *decreases* total revenue, demand is relatively inelastic.
- If the price cut *does not change* total revenue, demand is unit elastic.

The values in Table 4-1 are used to construct the demand and total revenue curves in Figure 4-3.

Table 4-1: Price, Demand, Total Revenue, and Elasticity

Price $	Quantity units	Total Revenue $	Elasticity
1	50	50	
			–0.16
2	45	90	
			–0.29
3	40	120	
			–0.47
4	35	140	
			–0.69
5	30	150	
			–1
6	25	150	
			–1.44
7	20	140	
			–2.14
8	15	120	
			–3.4
9	10	90	
			–6.33
10	5	50	

Note that we have used the arc elasticity formula to calculate the elasticities in this table.

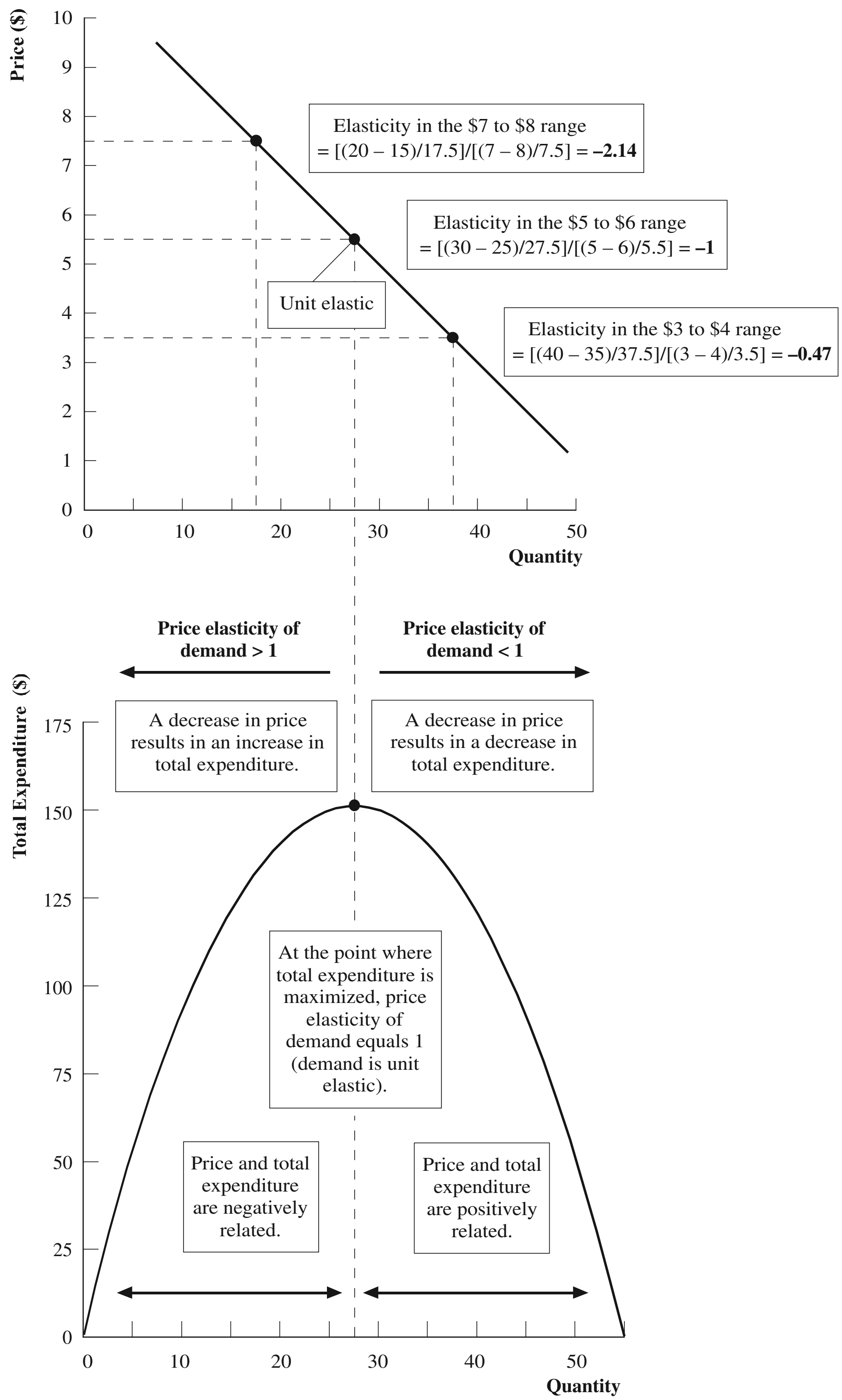
$|E_p| = 1 \Rightarrow$ TE remains the same regardless of the direction of change in price.

$|E_p| < 1 \Rightarrow$ If P↑, TE↑ If P↓, TE↓

$|E_p| > 1 \Rightarrow$ If P↑, TE↓ If P↓, TE↑

TE is maximized when $|E_p| = 1$

Figure 4-3: Elasticity and Total Expenditure



Total Revenue and Price Elasticity

Looking at things from a producer's perspective, the change in the total amount of money earned from sales also depends on sensitivity of quantity demanded to changes in price.

- If the demand curve facing a producer is relatively elastic, an *increase* in price will *decrease* total revenue.
- If the demand curve facing a producer is relatively inelastic, an *increase* in price will *increase* total revenue.
- If the demand curve facing a producer is unit elastic, an *increase* in price will not change total revenue.

Note that if a producer is currently charging a price that lies in the inelastic region of the demand curve, she can increase total revenue by increasing her prices. The benefit of higher prices would outweigh the negative impact of lower quantities sold. Further, since a lower quantity is being sold (and produced) total costs would also fall, which implies a certain boost in profitability. Therefore, no producer would knowingly set a price that falls in the inelastic region of the demand curve.

Income Elasticity of Demand

Income elasticity of demand measures the responsiveness of demand for a particular good to a change in income, holding all other things constant.

$$ED_I = \frac{\%\Delta QD_x}{\%\Delta I} = \frac{\Delta QD_x / QD_x}{\Delta I / I} = \left(\frac{\Delta QD_x}{\Delta I}\right)\left(\frac{I}{QD_x}\right) \quad \text{... (Equation 19)}$$

Same as coefficient on I in market demand function (Equation 11)

$$E_I = \frac{\%\text{ change in quantity demanded}}{\%\text{ change in income}}$$

Income elasticity of demand can be positive, negative, or zero. Products are classified along the following lines:

If income elasticity is *greater* than 1, demand is income elastic, and the product is classified as a normal good.

- As income *rises*, the percentage increase in demand *exceeds* the percentage change in income.
- As income *increases*, a consumer spends a *higher proportion* of her income on the product.

$E_I > 1 \Rightarrow$ Normal good (income elastic)

$0 < E_I < 1 \Rightarrow$ Normal good (income inelastic)

$E_I < 0 \Rightarrow$ Inferior good

If income elasticity lies *between* zero and 1, demand is income inelastic, but the product is still classified as a normal good.

- As income *rises*, the percentage *increase* in demand is *less* than the percentage increase in income.
- As income *increases*, a consumer spends a *lower proportion* of her income on the product.

If income elasticity is *less than zero (negative)*, the product is classified as an inferior good.

- As income *rises*, there is a *negative* change in demand.
- The *amount* spent on the good *decreases* as income *rises*.

For some goods, in some income ranges, income may have no impact on demand for the good. For these goods, income elasticity of demand equals 0.

Note that when income changes, there is a shift in the demand curve (change in demand). An increase in income results in an increase in demand (shift in demand curve to the right) for normal goods, and a decrease in demand (shift in demand curve to the left) for inferior goods. Further, an income elasticity of demand of 0.5 for a particular good means that whenever income increases by 1%, the quantity demanded **at each price** would rise by 0.5%.

Cross-Price Elasticity of Demand

Cross elasticity of demand measures the responsiveness of demand for a particular good to a change in price of *another* good, holding all other things constant.

Same as coefficient on P_Y in market demand function (Equation 11)

$$ED_{Py} = \frac{\%\Delta QD_x}{\%\Delta P_y} = \frac{\Delta QD_x / QD_x}{\Delta P_y / P_y} = \left(\frac{\Delta QD_x}{\Delta P_y}\right)\left(\frac{P_y}{QD_x}\right) \quad \text{... (Equation 20)}$$

$$E_C = \frac{\text{\% change in quantity demanded}}{\text{\% change in price of substitute or complement}}$$

Substitutes

If the price of Burger King®'s burgers were to go up, what would be the effect on demand for McDonald's® burgers?

$E_C > 0 \Rightarrow$ substitutes

$E_C < 0 \Rightarrow$ complements

For most people, these are close *substitutes* for each other. An increase in price of Burger King®'s burgers will result in a significant increase in demand for McDonald's® burgers as consumers switch to the relatively lower priced substitute.

The magnitude of the cross elasticity figure tells us how closely the two products serve as substitutes for each other. A *high* value indicates that the products are very *close* substitutes (i.e., if the price of one rises by only a small amount, demand for the other will rise significantly). For substitutes, the numerator and denominator of the cross-elasticity formula head in the *same* direction. Therefore cross-price elasticity of demand for substitutes is *positive.* Note that two products are classified as substitutes if the cross-price elasticity of demand is positive, regardless of whether they would actually be considered similar.

Complements

If the price of playing a round of golf on a golf course were to rise, what would be the effect on demand for golf balls?

Since playing a game of golf is impossible without golf balls, these products are *complements* for each other. An *increase* in price of using a golf course will reduce the number of rounds of golf played, and bring about a *decrease* in demand for golf balls.

For complements, the numerator and denominator of the cross elasticity formula head in *opposite* directions. Therefore, the cross elasticity of demand for complements is *negative.* Note that two products are classified as complements if the cross-price elasticity of demand is negative, regardless of whether they are typically consumed as a pair.

The *absolute value* of the cross-elasticity figure tells us how closely consumption of the two products is tied together and how closely they serve as complements for each other. A *high* absolute number indicates very *close* complements. If the price of one *rises,* consumers will *significantly reduce* their demand for the other.

Also note that for substitutes, an increase in the price of another good results in an increase in demand (shift in demand to the right), while for complements, an increase in price of another good results in a decrease in demand (shift in demand to the left).

Calculating Demand Elasticities from Demand Functions

Given Equations 17, 19, and 20, we can easily calculate the own-price, cross-price, and income elasticities of demand for gasoline. Recall that the market demand function (Equation 11) for gasoline was given as:

$$QD_G = 7{,}500 - 500P_G + 100I - 50P_A$$

Let's calculate the own-price, income, and cross-price elasticities of demand assuming that P_G = \$3/gallon, I = \$60,000/year, and P_A = \$30,000. The first step is to determine quantity demanded given the above values for the independent variables:

$QD_G = 7{,}500 - 500(3) + 100(60) - 50(30) = 10{,}500$ gallons. Given the quantity demanded, we calculate the different elasticities by simply plugging numbers into the elasticity formulas:

$$ED_{Px} = \frac{\%\Delta QD_x}{\%\Delta P_x} = \frac{\Delta QD_x / QD_x}{\Delta P_x / P_x} = (-500)\left(\frac{3}{10{,}500}\right) = -0.143$$

Coefficient on own-price in market demand function (Equation 11)

$$ED_I = \frac{\%\Delta QD_x}{\%\Delta I} = \frac{\Delta QD_x / QD_x}{\Delta I / I} = (100)\left(\frac{60}{10{,}500}\right) = 0.571$$

Same as coefficient on I in market demand function (Equation 11)

$$ED_{Py} = \frac{\%\Delta QD_x}{\%\Delta P_y} = \frac{\Delta QD_x / QD_x}{\Delta P_y / P_y} = (-50)\left(\frac{30}{10{,}500}\right) = -0.143$$

Same as coefficient on P_Y in market demand function (Equation 11)

- Since the absolute value of own-price elasticity lies between 0 and 1, we conclude that demand is relatively inelastic at a price of $3/gallon.
- Since income elasticity is positive and lies between 0 and 1, we conclude that gasoline is a normal good.
- Since cross-price elasticity is negative, we conclude that gasoline and automobiles are complements.

Reading 14: Demand And Supply Analysis: Consumer Demand

LESSON 1: INDIFFERENCE CURVES AND THE OPPORTUNITY SET

In this reading, we describe the foundations of demand analysis through consumer choice theory.

LOS 14a: Describe consumer choice theory and utility theory.

In describing utility theory, we use the concept of consumption bundles extensively. A consumption bundle is defined as a basket of goods and services that a consumer would like to consume. Two baskets of goods may each contain the exact same goods, but as long as the quantity of at least one good is different across the two bundles, they will be considered two distinct bundles.

Axioms of the Theory of Consumer Choice

Assumption of complete preferences: Given two bundles, a consumer is positively able to state which bundle she prefers to the other (if in fact she prefers one to the other) or that she is indifferent between the two. Even if the bundles are completely different and composed of entirely different goods, this assumption leaves no room for a response along the lines of "I just can't compare the two."

Assumption of transitive preferences: Given three bundles, A, B, and C, if a consumer prefers Bundle A to Bundle B, and prefers Bundle B to Bundle C, then it must be the case that she prefers Bundle A to Bundle C.

Assumption of nonsatiation: It can never be the case (for at least one of the goods in the bundle) that the consumer could at some stage have so much of the good that she would refuse more of the good even if it were free.

For some goods, more is actually worse (e.g., pollution). For such items, we define removal of the item as the good. For example, more removal of pollution would be preferred by the consumer.

The Utility Function

The utility function translates each bundle of goods and services into a single number (expressed in terms of utils, which are basically quantities of wellbeing), that allows us to rank the different bundles based on consumer preferences. In general a utility function can be represented as:

$$\boxed{U = f(Q_{x_1}, Q_{x_2}, \ldots, Q_{x_n})}$$

In this reading, we will primarily be working with a variety of different bundles that each contain only two goods (milk and chocolate) but in different quantities. Let's assume that:

- Bundle A contains 4 ounces of milk ($M_A = 4$) and 3 bars of chocolate ($C_A = 3$) for a total utility of 10 utils.
- Bundle B contains 5 ounces of milk ($M_B = 5$) and 1 bar of chocolate ($C_B = 1$) for a total utility of 12 utils.

Given our consumer's tastes and preferences, we can state that she prefers Bundle B to Bundle A. It is very important for us to note that utility functions offer an ordinal ranking, not a cardinal ranking. They allow us to determine which bundle is preferred but do not facilitate the calculation and ranking of differences between bundles.

LOS 14b: Describe the use of indifference curves, opportunity sets, and budget constraints in decision making.

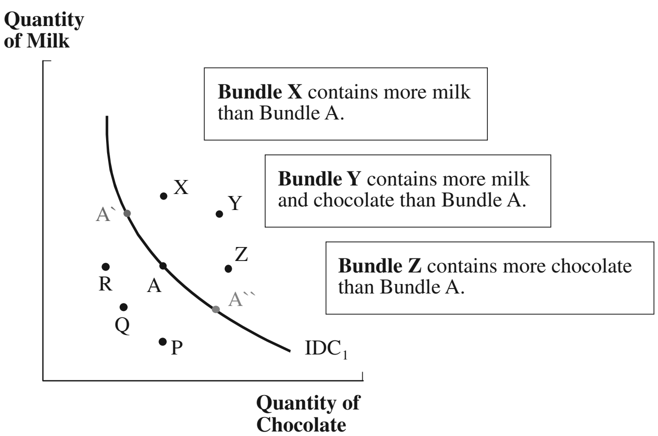
LOS 14c: Calculate and interpret a budget constraint.

Indifference Curves: The Graphical Portrayal of the Utility Function

We consider bundles with only two goods so that a two-dimensional graph can be used to illustrate different combinations.

Indifference curves represent all the bundles of two goods that yield exactly the same level of satisfaction for a consumer. Stated differently, an indifference curve represents all the combinations of two goods that the consumer is indifferent between.

Figure 1-1: An Indifference Curve



Let's start with Bundle A, which contains 4 ounces of milk and 3 bars of chocolate ($M_A = 4$; $C_A = 3$) and offers 10 utils of satisfaction to the consumer (see Figure 1-1). The first thing that we must understand when drawing up indifference curves is that due to the assumption of nonsatiation (i.e., more is always better):

- All points directly above Point A (e.g., Point X), directly to the right of Point A (e.g., Point Z), and above and to the right of Point A (e.g., Point Y) offer more utility than Bundle A. This set of bundles is known as the preferred-to-Bundle-A set.
- All points directly below Point A (e.g., Point P), directly to the left of Point A (e.g., Point R), and below and to the left of Point A (e.g., Point Q) offer less utility than Bundle A. This set of bundles is known as the dominated-by-Bundle-A set.

Now assume that there are two more bundles, Bundle A` and Bundle A`` that are also entirely composed of milk and chocolate, and offer the consumer the same total utility as Bundle A (i.e., 10 utils).

- Bundle A` consists of more milk but less chocolate than Bundle A. Therefore, it lies above and to the left of Bundle A.
- Bundle A`` consists of more chocolate but less milk than Bundle A. Therefore, it lies below and to the right of Bundle A.

When we plot all such combinations of milk and chocolate that yield the same level of satisfaction to the consumer as Bundles A, A`, and A`` (10 utils) we obtain an indifference curve. The indifference curve, IDC_1 (in Figure 1-1), represents all combinations of milk and chocolate that offer 10 utils of satisfaction to the consumer. Now notice that:

- The preferred-to-Bundle-A set now consists of all bundles that lie above and to the right of IDC_1.

- The dominated-by-Bundle-A set now consists of all bundles that lie below and to the left of IDC_1.

Notice that the indifference curve has a negative slope. This implies that both milk and chocolate offer positive utility to the consumer. In order to maintain indifference between two different bundles, a decrease in the quantity of milk should be compensated by an increase in the amount of chocolate (and vice versa).

Think of the marginal rate of substitution of X for Y (rate of sacrificing Y to obtain more X), MRS_{XY}, as the negative of the slope of the indifference curve.

The indifference curve has a (negative) slope that equals $-\Delta Y/\Delta X$. As a result, MRS_{XY} equals $\Delta Y/\Delta X$.

If the slope of the indifference curve equals −3, it means that MRS_{XY} equals 3 or that the consumer would be willing to sacrifice 3 units of Good Y to obtain one more unit of Good X.

Also notice that the curve is convex when viewed from the origin (its slope becomes less steep as we move to the right). The slope of the curve represents the marginal rate of substitution of chocolate for milk (MRS_{CM}) (i.e., the quantity of milk that the consumer is willing to give up or sacrifice in order to obtain one additional unit of chocolate, holding utility constant).

Also notice that at Point A` the slope of the indifference curve is relatively steep. At this point, the consumer is willing to give up a considerable amount of milk to obtain more chocolate. On the other hand, at Point A`` the slope of the indifference curve is less steep. At this point, the consumer is willing to give up a relatively lower amount of milk to obtain more chocolate. This implies that the value the consumer places on chocolate (in terms of the amount of milk she is willing to give up) diminishes the more chocolate and the less milk she has (as we move rightward/downward along the indifference curve).

Indifference Curve Maps

An indifference curve map represents a consumer's entire set of indifference curves, where each indifference curve offers the consumer a different level of utility. The higher/more rightward an indifference curve lies, the greater the level of utility its representative bundles offer.

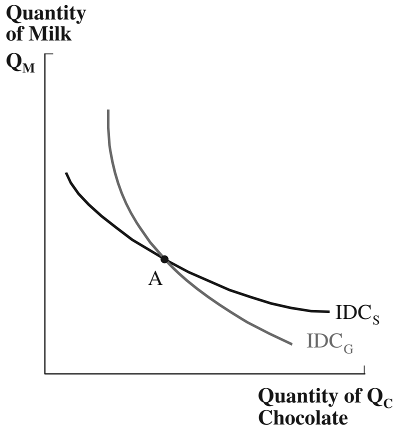
- Due to the completeness assumption (i.e., all available bundles can be compared), there must be one indifference curve that passes through every point.
- Due to the transitivity assumption, indifference curves for a particular consumer can never intersect each other.

Therefore, indifference curves are generally convex and negatively sloped. Further, they cannot cross.

Gains from Voluntary Exchange: Creating Wealth Through Trade

In the previous section, we mentioned that it is impossible for two indifference curves for the same consumer to ever cross. However, there is no such requirement for indifference curves for two consumers with different tastes and preferences. Figure 1-2 shows indifference curves for George (IDC_G) and Sally (IDC_S). Assume that they both are given identical quantities of milk and chocolate (Bundle A).

Figure 1-2: Indifference Curves for George and Sally



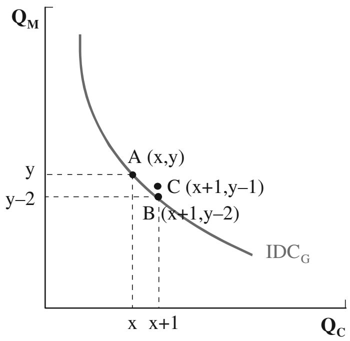
Note that since their indifference curves intersect at Point A, their slopes at this point are different.

- At Point A, IDC_G has a steeper slope than IDC_S, which means that George's marginal rate of substitution of chocolate for milk (MRS_{CM}) is greater than Sally's. Stated differently, George is willing to give up more milk to obtain an additional unit of chocolate compared to Sally. MRS_{CM} is measured as the negative of the slope of the indifference curve.
- This also implies that at Point A, Sally's marginal rate of substitution of milk for chocolate (MRS_{MC}) is greater than George's (i.e., she is willing to give up more chocolate to obtain an additional unit of milk compared to George). MRS_{MC} is measured as the negative of the inverse of the slope of the indifference curve.
- Therefore, at Point A, George has a relatively strong preference for chocolate, while Sally has a relatively strong preference for milk.

Suppose that the slope of George's indifference curve at Point A is –2, while that of Sally's indifference curve is –0.5. This means that George is willing to give up 2 units of milk to obtain 1 more unit of chocolate (George's $MRS_{CM} = 2$), while Sally is only willing to give up half a unit of milk to obtain 1 more unit of chocolate (Sally's $MRS_{CM} = 0.5$). Note that this also implies that Sally would give up 2 units of chocolate to obtain 1 more unit of milk (Sally's $MRS_{MC} = 2$). If they were able to exchange 1 unit of chocolate for 1 unit of milk, they would both be better off:

- George was willing to give up 2 units of milk to obtain 1 more unit of chocolate. The exchange enables him to increase his chocolate consumption by 1 unit and only give up 1 unit of milk. Relative to his original indifference curve, George is able to consume 1 extra unit of milk, which takes him to a higher indifference curve (see Figure 1-3 for detailed analysis).
- Sally was willing to give up 2 units of chocolate to obtain 1 more unit of milk. The exchange enables her to increase her milk consumption by 1 unit and only give up 1 unit of chocolate. Relative to her original indifference curve, Sally is able to consume 1 more unit of chocolate, which takes her to a higher indifference curve as well.

Figure 1-3: The Benefits of Voluntary Exchange



- At Point A, George consumes x units of chocolate and y units of milk.
- The slope of his IDC at Point A (slope = −2, MRS_{CM} = 2) implies that he is willing to give up 2 units of milk to obtain 1 more unit of chocolate.
- At Point B (on the same indifference curve) he consumes x+1 units of chocolate and y−2 units of milk.
- However, through a 1-for-1 exchange with Sally, he is able to move from Point A to Point C, where he consumes y−1 units of milk (as he gives one unit of milk to Sally) and x+1 units of chocolate (as he gets one unit of chocolate from Sally in return).
- Relative to Point B (which lies on his original IDC), at Point C, George consumes 1 more unit of milk (and the same quantity of chocolate), which means that Point C must lie on a higher indifference curve.

Conclusion: The voluntary exchange makes George (and Sally) better off as it allows them to move to consumption baskets that lie on higher (than their current) indifference curves.

Since both our consumers end up on higher indifference curves as a result of the exchange, there is greater total utility after the exchange. Note that after the exchange, George ends up with more chocolate and less milk than before (comparing Point A to Point C), while Sally ends up with more milk and less chocolate than before.

- As George gives up units of milk for more chocolate (moves rightward along the indifference curve), his MRS_{CM} declines (his indifference curve becomes less steep).
- As Sally gives up units of chocolate for more milk (moves leftward along the indifference curve), her MRS_{CM} increases (her indifference curve becomes steeper).
- As they continue to trade, eventually their MRS_{CM}'s reach equality. At this point, there are no longer any benefits from exchange.

The Opportunity Set: Consumption, Production, and Investment Choice

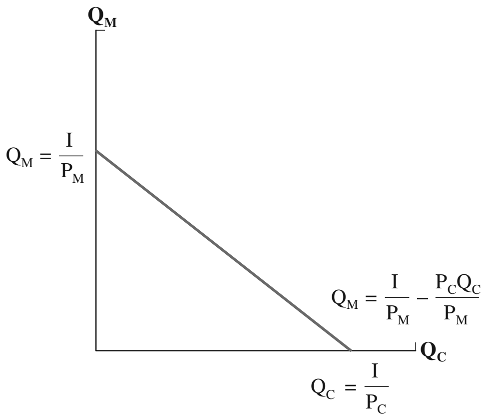
So far, we have examined the choices and tradeoffs that consumers are *willing* to make. Now we move into what dictates the choices and tradeoffs that consumers are actually *able* to make. Generally, consumers have limited incomes to purchase goods and services.

The Budget Constraint

George's preferences for milk versus chocolate were illustrated by IDC_G in the previous section. Now we define his income as I, the price of milk as P_M, and the price of chocolate as P_C. Since he cannot spend more on milk and chocolate than his total income per time period, and we assume that he has no reason not to spend all his income, his income constraint or budget constraint can be captured by the following expression, illustrated in Figure 1-4:

$$\boxed{P_M Q_M + P_C Q_C = I} \quad \text{... (Equation 1)}$$

Figure 1-4: The Budget Constraint



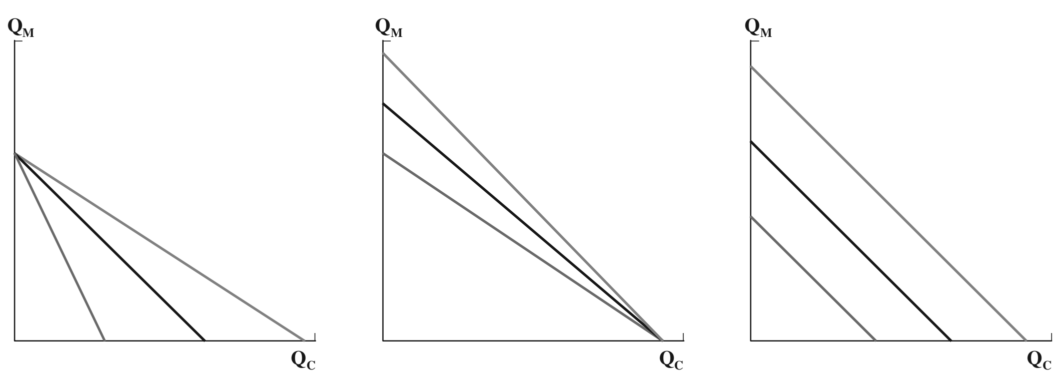
Given this budget constraint, the following should be fairly easy to understand:

- If George were to spend his entire income on milk, he would be able to purchase I/P_M units of milk.
- If George were to spend his entire income on chocolate, he would be able to purchase I/P_C units of chocolate.
- Manipulating Equation 1, his budget constraint can be defined by the following equation:

$$\boxed{Q_M = \frac{I}{P_M} - \frac{P_C Q_C}{P_M}} \qquad \text{slope equals} - \frac{P_C}{P_M}$$

Figure 1-5: Changing Prices and Income

1-5a: Change in P_C **1-5b: Change in P_M** **1-5c: Change in P_C and P_M**



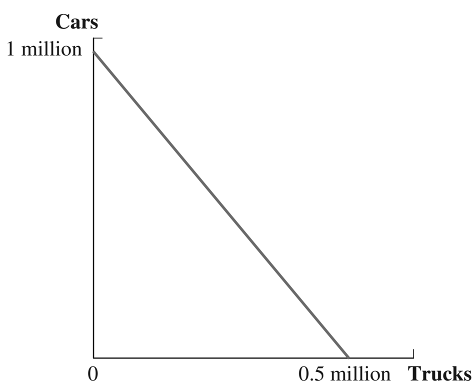
Note the following important points:

- The slope of the budget constraint equals $-P_C/P_M$. The slope basically identifies the quantity of milk that George would have to give up in order to purchase 1 more unit of chocolate.
- If the price of chocolate were to rise (fall), the horizontal intercept of the budget constraint would decrease (increase) and the budget constraint would look like the green (grey) line in Figure 1-5a.
- If the price of milk were to rise (fall), the vertical intercept (y-intercept) of the budget constraint would decrease (increase) and the budget constraint would look like the green (grey) line in Figure 1-5b.
- If the prices of both milk and chocolate were to increase (decrease) both the intercepts would decrease (increase) and the budget constraint would look like the green (grey) line in Figure 1-5c.

The Production Opportunity Set

Just like a consumer's income places a limit on the quantities of two goods she can consume (budget constraint), a producer's production capacity places a limit on her ability to produce two goods. For example, an automobile company's capacity limits the quantities of cars and trucks that it can supply over a period of time. If it wants to produce more cars, it must reduce its production of trucks and vice versa. A company's production possibility frontier (PPF) (Figure 1-6) shows the maximum units of one good that can be produced at each level of production for the other good.

Figure 1-6: The Production Possibility Frontier



From Figure 1-6, notice the following:

- If the entire production facility is devoted to the production of cars, the company would produce 1 million cars.
- If the entire production facility is devoted to the production of trucks, the company would produce 0.5 million trucks.
- For every additional truck that the company wants to produce, it must forego the production of 2 cars. The opportunity cost of producing 1 truck is 2 cars (negative of slope of PPF).
- For every additional car that the company wants to produce, it must forego the production of 0.5 trucks. The opportunity cost of producing 1 car is 0.5 trucks (negative of inverse of slope of PPF).

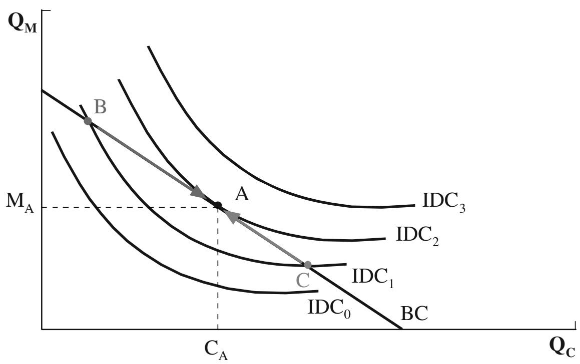
- In this analysis of the PPF, we have assumed that the opportunity cost of cars is constant, irrespective of the current production of cars. Similarly, the opportunity cost of trucks is constant, irrespective of the current production of trucks. As a result, our PPF is linear.
- A more realistic representation of the PPF would be a concave curve, with the marginal opportunity cost of producing trucks (in terms of production of cars sacrificed) increasing as more and more resources that specialize in car production are diverted to production of trucks.

Note: At this juncture, the curriculum moves into a brief discussion of the investment opportunity set. We have discussed this in detail in the readings on portfolio management.

LESSON 2: CONSUMER EQUILIBRIUM

LOS 14d: Determine a consumer's equilibrium bundle of goods based on utility analysis.

Figure 2-1: George's Indifference Curve Map and Budget Constraint



Recall that the slope of the budget constraint equals $-P_C/P_M$ and that the slope of the indifference curve equals $-MRS_{CM}$.

Figure 2-1 illustrates George's indifference curve map along with his budget constraint. The aim is to maximize utility given the prices of milk and chocolate, and his income. Therefore, George strives to attain the indifference curve that lies farthest from the origin without violating his budget constraint. The point of maximum affordable satisfaction occurs at the point of tangency (Point A) between his highest indifference curve (IDC_2) and budget constraint (BC). Note that at this point:

- The slope of the budget constraint (BC) equals the slope of the indifference curve (IDC_2).
- Consequently, the ratio of the price of chocolate to the price of milk (P_C/P_M) must equal the marginal rate of substitution of chocolate for milk (MRS_{CM}) at this point.
- MRS_{CM} is the rate at which the consumer is **willing** to sacrifice milk for chocolate. Further, the price ratio indicates the rate at which she **must** sacrifice milk for chocolate.
- Therefore, at equilibrium, the consumer is just willing to pay the opportunity cost that he must pay to obtain more chocolate.

In contrast, notice that at Point B:

- MRS_{CM} is greater than the price ratio. (The slope of the indifference curve is steeper than the slope of the budget line.)
- This implies that he is willing to give up milk to obtain chocolate at a rate that is greater than he must (given by the price ratio). In other words, he is willing to pay a higher price for chocolate (in terms of milk sacrificed) than he must until his chocolate consumption increases to C_A.
- Therefore, he is better off moving downward along his budget constraint until he reaches Point A.

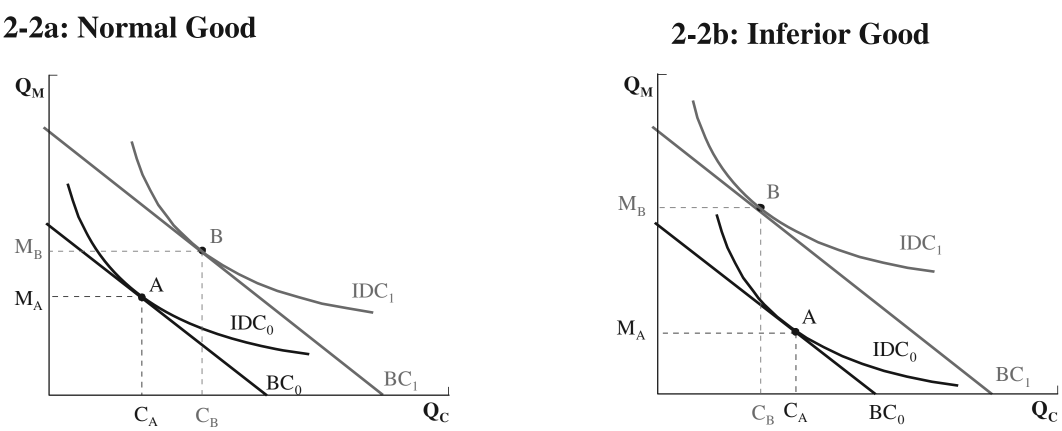
While at Point C:

- MRS_{CM} is lower than the price ratio. (The slope of the indifference curve is flatter than the slope of the budget line.)
- This implies that he is willing to give up milk to obtain chocolate at a rate that is lower than he must (given by the price ratio). In other words, he is currently paying a higher price for chocolate (in terms of milk sacrificed) than he is willing to until his chocolate consumption falls to C_A.
- Therefore, he is better off moving upward along his budget constraint until he reaches Point A.

Consumer Response to Changes in Income: Normal and Inferior Goods

We learned previously that a consumer's income and the prices of goods and services place a constraint on his consumption behavior. A change in income or product prices leads to a change in the point of tangency between the budget constraint and the highest attainable indifference curve, and therefore a change in consumption behaviour.

Figure 2-2: Changes in Income



Case 1: An Increase in Income

An increase in income results in a parallel, outward (to the right) shift in the budget constraint. If both goods are normal goods, when his income increases, the individual will increase his consumption of both goods (see Figure 2-2a).

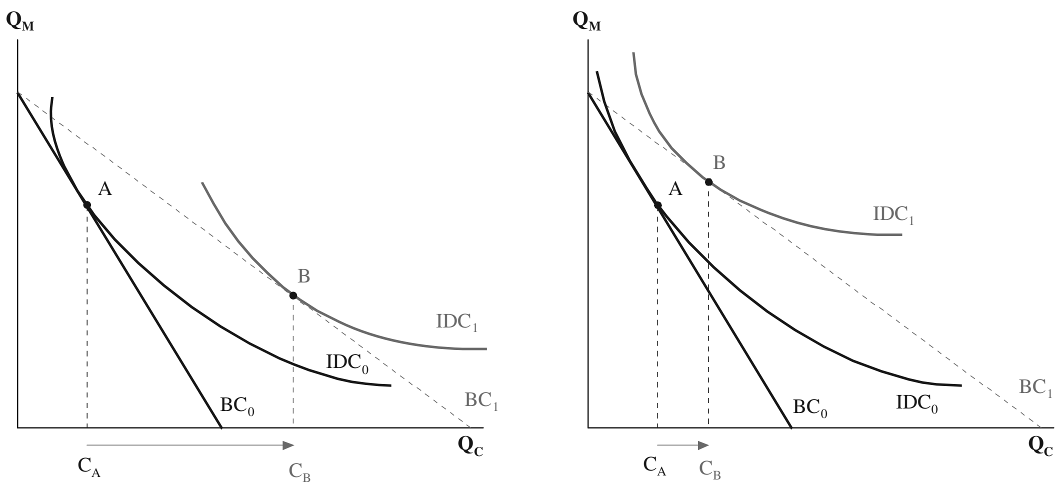
If one good is a normal good while the other is an inferior good, in response to an increase in income the individual will increase his consumption of the normal good, but decrease his consumption of the inferior good (see Figure 2-2b). We assume that chocolate is the inferior good and milk is the normal good. Notice that as the consumer's budget constraint moves to the right from BC_0 to BC_1 (as income increases) he moves from IDC_0 to IDC_1 and his consumption basket moves from Point A to Point B. Note that Basket B contains less of the inferior good, chocolate (C_B versus C_A), and more of the normal good, milk (M_B versus M_A).

A normal good is a good with a positive income elasticity of demand. An increase in income leads to an increase in quantity consumed.

An inferior good is a good with a negative income elasticity of demand. An increase in income leads to a decrease in quantity consumed.

Figure 2-3: Elastic and Inelastic Responses to a Decrease in Price

2-3a: Elastic Response **2-3b: Inelastic Response**



Case 2: Changes in Price

Now we assume that the price of milk remains constant and that of chocolate decreases. A decrease in the price of chocolate results in an increase in the horizontal intercept (x-intercept) of the budget constraint, while the vertical intercept (y-intercept) remains the same. Figures 2-3a and 2-3b illustrate two types of responses that the consumer may have to this decrease in the price of chocolate.

1. The quantity of chocolate consumed may increase by a relatively significant amount, which would imply that demand for chocolate is relatively elastic.
2. The quantity of chocolate consumed may increase by a relatively small amount, which would imply that demand for chocolate is relatively inelastic.

Consumer's Demand Curve from Preferences and Budget Constraints

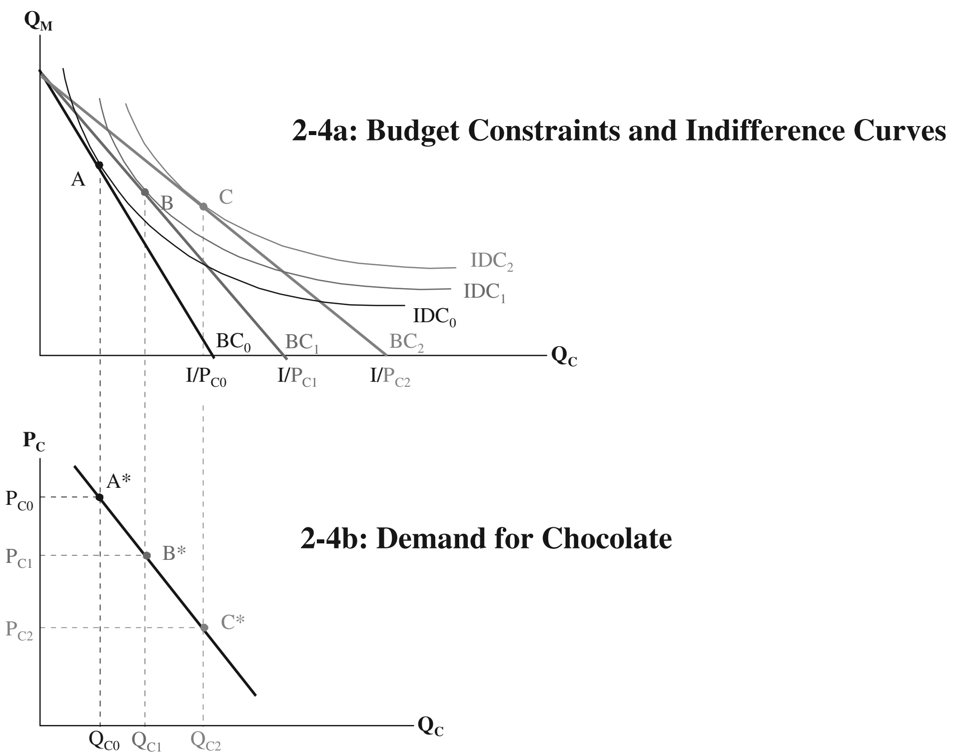
We use Figure 2-4 to illustrate the derivation of the demand curve for chocolate. Note that in deriving the demand curve, we assume that the independent variable is own-price (price of chocolate), while income and prices of other goods (price of milk) are held constant.

Figure 2-4a has the quantity of milk on the vertical axis and the quantity of chocolate on the horizontal axis. It also contains the consumer's indifference curve map along with an initial budget constraint (BC_0) which is drawn up assuming that the consumer's income equals I and the price of chocolate and milk are P_{C0} and P_{M0} respectively.

Figure 2-4b has the price of chocolate (P_C) on the vertical axis and the quantity of chocolate (Q_C) on the horizontal axis. Note that the consumer's income (I) and the price of milk (P_M) are held constant throughout the analysis that follows.

- With the price of chocolate at P_{C0}, the budget constraint, B_{C0}, is tangent to the highest indifference curve (IDC_0) at Point A. At this point, the quantity of chocolate in the consumption basket equals Q_{C0}. Therefore, on Figure 2-4b we plot Point A* with coordinates Q_{C0}, P_{C0}.
- If the price of chocolate decreases to P_{C1}, the budget constraint pivots outward as the horizontal intercept of the constraint increases. The point of intersection between the new budget constraint, BC_1, and the highest indifference curve, IDC_1, occurs at Point B. At this point, the quantity of chocolate in the consumption basket equals Q_{C1}. Therefore, on Figure 2-4b we plot Point B* with coordinates Q_{C1}, P_{C1}.
- If the price of chocolate decreases further to P_{C2}, the budget constraint pivots further outward as the horizontal intercept of the constraint increases even more. The point of intersection between the new budget constraint, BC_2, and the highest indifference curve, IDC_2, occurs at Point C. At this point, the quantity of chocolate in the consumption basket equals Q_{C2}. Therefore, on Figure 2-4b we plot Point C* with coordinates Q_{C2}, P_{C2}.
  These three price-quantity combinations (Points A*, B*, and C*) trace out our consumer's demand curve for chocolate.

Figure 2-4: Deriving the Demand Curve



LESSON 3: INCOME AND SUBSTITUTION EFFECTS

LOS 14e: Compare substitution and income effects.

LOS 14f: Distinguish between normal goods and inferior goods and explain Giffen goods and Veblen goods in this context.

Substitution and Income Effects for a Normal Good

The law of demand states that when the price of a good falls (rises), quantity demanded increases (decreases). There are two main reasons for this:

1. The substitution effect: The good becomes relatively cheaper compared to other goods, so more of the good gets substituted for other goods in the consumer's consumption basket.
2. The income effect: The consumer's real income increases (in terms of the quantity of goods and services that can be purchased with the same dollar income). If the good is a normal good (which most goods are), the increase in real income (due to the decrease in its price) leads to an increase in quantity purchased.

The reduction in the price of chocolate is a good thing for the consumer. The idea behind moving BC_1 inward until it is parallel to the original indifference curve (IDC_0) is that we want to offset the benefit of the decrease in price by reducing real income and, at the same time, ensure that the consumer remains as well off as before the price change (moves along his original indifference curve). Once the real income effect has been isolated and removed, what remains is the response resulting from the substitution effect alone.

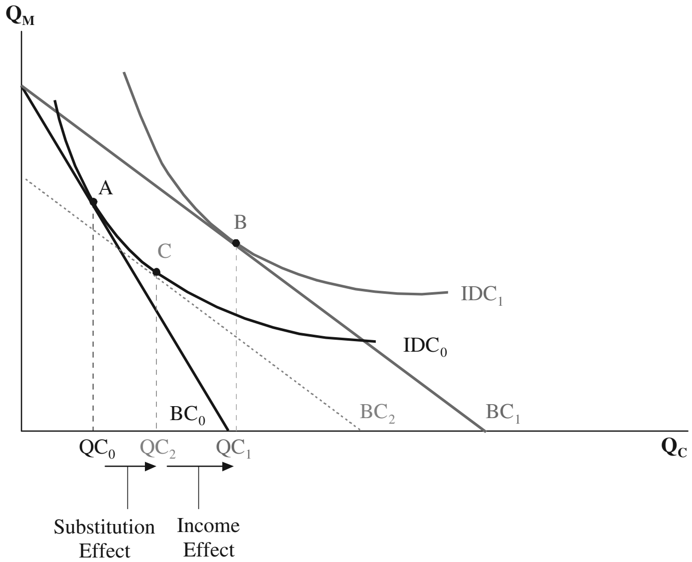
The analysis that follows demonstrates how these two effects on quantity demanded can be separated.

Figure 3-1 shows the consumer's original budget constraint, BC_0, and his highest attainable indifference curve (IDC_0) given the budget constraint. The point of tangency between the two (Point A) indicates the quantities of milk (QM_0) and chocolate (DC_0) that the consumer will purchase given their prices (PM_0 and PC_0 respectively) and his income (I).

A decrease in the price of chocolate (from PC_0 to PC_1) results in the budget constraint pivoting (from BC_0 to BC_1). Given this new budget constraint, the highest attainable indifference curve is now IDC_1, and the point of tangency between IDC_1 and BC_1 results in a new optimal consumption basket defined by Point B. The quantity of chocolate purchased has increased from QC_0 to QC_1, but this change in quantity includes the impact of both the income and substitution effects of the price change. In order to isolate the two effects:

- We drag the new budget constraint (BC_1) inward (toward the origin) until it is parallel to the original indifference curve (IDC_0). This new, dragged-back budget constraint (BC_2) is tangent to the original indifference curve (IDC_0) at Point C, where the quantity of chocolate purchased equals QC_2.
- The movement from Point A to Point C is caused by the substitution effect (movement along the same original indifference curve) of the price change. The substitution effect results in an increase in quantity of chocolate consumed from QC_0 to QC_2.
- The movement from Point C to Point B is caused by the income effect of the price change. The income effect results in an increase in quantity of chocolate consumed from QC_2 to QC_1.

Figure 3-1: Substitution and Income Effects for a Normal Good



The substitution effect always goes in the opposite direction of the price change. In our analysis above, as the price of chocolate fell the substitution effect resulted in an increase in quantity demanded (from QC_0 to QC_2). The substitution effect is the result from changing the budget constraint from BC_0 to BC_2, and moving along the original indifference curve, IDC_0, while maintaining tangency. Since the price ratio (negative of slope of the budget line) has fallen from PC_0/PM_0 to PC_1/PM_0 (as PC_1 is less than PC_0) the consumer moves rightward along his indifference curve, sacrificing milk for chocolate until his MRS_{CM} falls to a level that equals the new price ratio, PC_1/PM_0.

Sellers try to isolate the income and substitution effects when they analyze demand for their products and try to come up with creative pricing schemes (e.g., two-part tariff pricing) to extract consumer surplus from customers. Example 3-1 illustrates this.

Example 3-1: Two-Part Tariff Pricing

Max Lin's monthly demand for visits to the gym (QD_G) is given by $QD_G = 25 - 5P_G$ where P_G refers to price per visit. The price charged by the gym for each visit equals its marginal costs of a visit ($2).

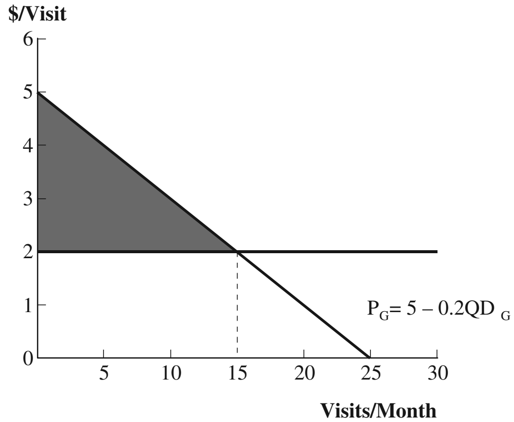
1. How many times would Max visit the gym each month if the gym were to charge $2 per visit?
2. Calculate Max's consumer surplus if the gym were to charge $2 per visit.
3. Explain how the gym could benefit by charging Max a monthly membership fee in addition to $2/visit.

Solution

Max's demand curve is given by the following inverse demand function:

$$P_G = 5 - 0.2QD_G$$

Figure 3-2: Max's Demand Curve for Gym Visits per Month



1. At a price of \$2 per visit, Max would visit the gym 15 times a month.
2. His consumer surplus equals the area of the triangle colored in green on Figure 3-2: $0.5 \times 15 \times (5 - 2) = \$22.50$
3. In addition to charging the \$2/visit fee, the gym could capture Max's entire consumer surplus by charging him a fixed monthly membership fee amounting to \$22.50. This kind of pricing mechanism is known as a two-part tariff because it has a price/unit component based on the quantity purchased plus a fixed-fee (that aims to capture any consumer surplus given the per-unit price).

Income and Substitution Effects for an Inferior Good

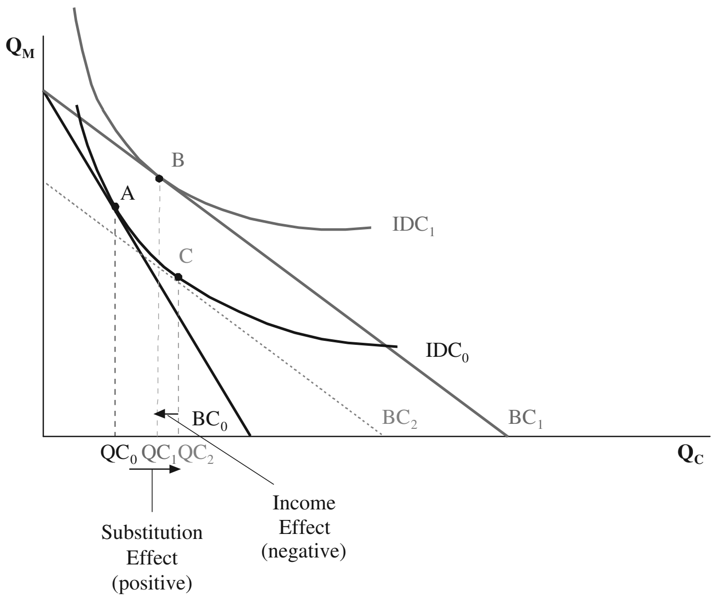
An inferior good is one that has negative income elasticity of demand. When income increases (decreases) demand for these goods falls (rises). When price falls, demand for inferior goods is also influenced by the income and substitution effects but, as you will see in the analysis that follows, the two effects drag quantity demanded in opposite directions.

Recall that the substitution effect is measured along the original indifference curve.

We assume that chocolate is the inferior good in this analysis (see Figure 3-3). A decrease in the price of chocolate pivots the budget constraint to BC_1 and the quantity of chocolate demanded increases to QC_1 as the consumer moves to a higher indifference curve, IDC_1. This increase in quantity demanded (from QC_0 to QC_1) incorporates the impact of both the income and substitution effects of the price change. To isolate the two, we drag the new budget constraint, BC_1, toward the origin until it is tangent to the original indifference curve, IDC_0. The point of intersection between the "dragged-back" budget constraint, BC_2, and the original indifference curve occurs at Point C, where the quantity of chocolate equals QC_2. Notice that QC_2 is greater than QC_1, which implies that the substitution effect alone ($QC_2 - QC_0$) actually exceeds the combined impact of the income and substitution effects ($QC_1 - QC_0$). The implication here is that the income effect is negative.

The decrease in price of chocolate increases the consumer's real income, but because chocolate is an inferior good, the income effect is negative. The overall impact on quantity demanded of the decrease in price is still positive (quantity demanded increases from QC_0 to QC_1), but the change in quantity demanded is not very significant as the negative income effect partially offsets the positive substitution effect of the price reduction. Demand for chocolate (when it is assumed to be an inferior good) is therefore less elastic than if the two effects reinforced each other (as was the case in the analysis in the previous section when chocolate was assumed to be a normal good).

Figure 3-3: Substitution and Income Effects for an Inferior Good



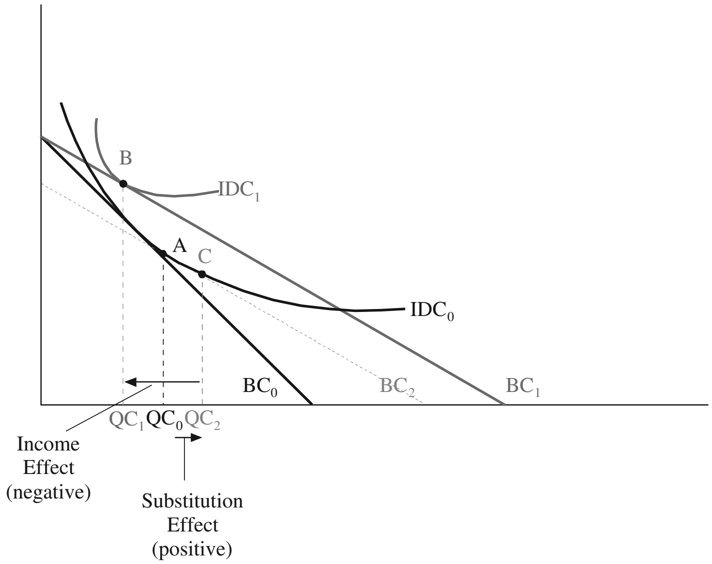
Giffen Goods

A Giffen good is a special case of an inferior good where the negative income effect of a decrease in price of the good is so strong that it outweighs the positive substitution effect. Therefore, for a Giffen good, quantity demanded actually falls when there is a decrease in price, which makes the demand curve upward sloping. The analysis that follows illustrates the income and substitution effects for a Giffen good.

We assume that chocolate is the Giffen good. A decrease in the price of chocolate pivots the budget constraint to BC_1 (see Figure 3-4). With this new budget constraint, the highest indifference curve (IDC_1) is tangent to the constraint at Point B, where quantity demanded is actually less than at Point A (QC_1 is less than QC_0). When we drag the new budget constraint inwards toward the original budget constraint, we see that BC_2 is tangent to the original indifference curve at Point C where quantity demanded equals QC_2. Since QC_2 is greater than QC_0 we conclude that the price decrease has a positive substitution effect on the quantity of chocolate ($QC_2 - QC_0$), but we also know that the overall impact of both the income and substitution effects on quantity demanded is negative ($QC_0 - QC_1$). The only logical explanation for this is that chocolate is a Giffen good and the negative income effect of the price reduction is so significant that it outweighs the positive substitution effect, and results in an overall decrease in quantity demanded even though the price of chocolate fell.

Note that all Giffen goods are inferior goods, but not all inferior goods are Giffen goods.

Figure 3-4: Giffen Goods



Veblen Goods

Sometimes, the price tag of a good itself determines its desirability for consumers. For example, with status goods such as expensive jewelery, the high price itself adds to the utility from the good, such that the consumer values the item more if it has a higher price. Such goods are known as Veblen goods, and it is argued that their demand curves are also upward sloping just like Giffen goods. However, there remains a fundamental difference between the two.

- Giffen goods **are inferior goods**. They are not status goods. An increase in income would reduce demand for them (due to negative income elasticity of demand).
- Veblen goods **are not inferior goods**. An increase in income would not lead to a decrease in demand.

READING 15: DEMAND AND SUPPLY ANALYSIS: THE FIRM

LESSON 1: TYPES OF PROFIT MEASURES

Objectives of the Firm

In this reading we assume that the primary objective of the firm is to maximize profits. Generally speaking, profits are calculated as the difference between total revenue and total costs. Total revenue is a function of the selling price and quantity sold, which are both determined by demand and supply in the market for the good produced by the firm. Costs are a function of the level of output, the efficiency of the firm's production processes and resource prices (which depend on demand and supply in resource markets).

LOS 15a: Calculate, interpret, and compare accounting profit, economic profit, normal profit, and economic rent.

Types of Profit Measures

In this reading, whenever we talk about profits, we are referring to economic profits unless we explicitly state that the discussion revolves around accounting profits.

Accounting Profit

Accounting profit (also known as net profit, net income, and net earnings) equals revenue less all accounting (or explicit) costs. Accounting costs are payments to nonowner parties for goods and services supplied to the firm and do not necessarily require a cash outlay.

Accounting profit (loss) = Total revenue − Total accounting costs.

Economic Profit and Normal Profit

Economic profit (also known as abnormal profit or supernormal profit) is calculated as revenue less all economic costs (economic costs include explicit and implicit costs). Alternatively, economic profit can be calculated as accounting profit less all implicit opportunity costs that are not included in total accounting costs.

Economic profit = Total revenue − Total economic costs

Economic profit = Total revenue − (Explicit costs + Implicit costs)

Economic profit = Accounting profit − Total implicit opportunity costs

Example 1-1: Economic Profit, Accounting Profit, and Normal Profit

Consider two companies, a startup, SU, and a public limited company, PLC. The following table includes revenue and cost information for the two companies for $:

	Start-up (SU) $	Public Limited Company (PLC) $
Total revenue	4,500,000	60,000,000
Total accounting costs	4,100,000	57,000,000

The owner of SU takes a salary reduction of $50,000 relative to the job he gave up to work on the start-up. He also invested $2,000,000 in the business on which he could expect to earn $350,000 annually if he had invested the money in a similar-risk investment. PLC has equity investment worth $25,000,000 on which shareholders require a return of 7%.

Given that there are no other identifiable implicit opportunity costs for either firm, calculate accounting and economic profits for both the companies for 2011.

Solution

Accounting profit (loss) = Total revenue – Total accounting costs

Accounting profit (SU) = $4,500,000 – $4,100,000 = $400,000
Accounting profit (PLC) = $60,000,000 – $57,000,000 = $3,000,000

Economic profit (loss) = Accounting profit (loss) – Total implicit opportunity costs

The salary cut taken by SU's owner relative to his previous job ($50,000) and the investment income foregone on the money invested in SU ($350,000) are both implicit opportunity costs that must be deducted from SU's accounting profit to determine its economic profit.

Economic profit (SU) = $400,000 – ($50,000 + $350,000) = 0

The cost of equity capital is an implicit opportunity cost for PLC. Therefore, economic profit is calculated as:

Economic profit (PLC) = $3,000,000 – ($25,000,000 × 0.07) = $1,250,000

Notice that for SU, economic profit equals 0 as total revenues and total economic costs are equal. Since it just meets its opportunity costs, we can state that SU is earning normal profit of $400,000. Normal profit is the level of accounting profit that is required to cover the implicit opportunity costs that are not included in accounting costs. PLC's normal profit equals $1,750,000 (dollar cost of equity).

Accounting profit, economic profit, and normal profit are linked by the following equation:

Accounting profit = Economic profit + Normal profit

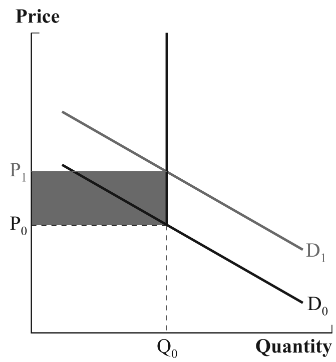
Notice that:

- When accounting profit equals normal profit, economic profit equals 0 (as is the case with SU)
- When accounting profit is greater than normal profit, economic profit is positive (as is the case with PLC).
- When accounting profit is less than normal profit, economic profit is negative. The firm incurs an economic loss.

Economic Rent

Economic rent can be defined as the payment for a good or service beyond the minimum amount needed to sustain supply. Economic rent results when the supply of a good is fixed (i.e., the supply curve is vertical or perfectly inelastic and the market price is higher than the minimum price required to bring the good onto the market). Since supply, S_0, is perfectly inelastic (see Figure 1-1), the demand curve determines the price and the level of economic rent. Initially, demand is at D_0, so price equals P_0. We assume that P_0 is the minimum price required by the business as it results in a normal profit for the firm. When demand increases to D_1, price increases to P_1 and since P_1 is greater than P_0 (the minimum price required by the firm), the firm now makes economic rent equal to the area of the rectangle shaded in green: $(P_1 - P_0) \times Q_0$.

Figure 1-1: Economic Rent



Economic rent is therefore, created in markets (e.g., land and certain commodities) where supply is tight (because of nature or government policy) such that when price increases, quantity supplied does not change (or does not change by much). The concept of economic rent can also be applied in the investment management arena. Investors look to put money in companies whose products are expected to witness an increase in demand in the future and have relatively inelastic supply. When demand does rise, prices would increase dramatically, resulting in an increase in shareholder wealth.

Comparison of Profit Measures

- In the short run, the normal profit rate is relatively stable, which makes accounting and economic profits the variable items in the profit equation.
- Over the long run, all three types of profit are variable.
- A firm must make at least a normal profit to stay in business in the long run.

Table 1-1 compares the three profit measures in terms of their impact on shareholder wealth.

Table 1-1: Relationship of Accounting, Normal, and Economic Profit to Equity Value[1]

Relationship Between Accounting Profit and Normal Profit	Economic Profit	Firm's Market Value of Equity
Accounting profit > Normal profit	Economic profit > 0 and firm is able to protect economic profit over the long run	Positive effect
Accounting profit = Normal profit	Economic profit = 0	No effect
Accounting profit < Normal profit	Economic profit < 0 implies economic loss	Negative effect

1 - Exhibit 2, Volume 2, CFA Program Curriculum 2015

LESSON 2: ANALYSIS OF REVENUE AND COSTS

LOS 15b: Calculate and interpret and compare total, average, and marginal revenue.

Total, Average, and Marginal Revenue

Table 2-1 lists the formulas for calculating different revenue items:

Table 2-1: Summary of Revenue Terms[2]

Revenue	Calculation
Total revenue (TR)	Price times quantity (P × Q), or the sum of individual units sold times their respective prices; Σ(Pi × Qi)
Average revenue (AR)	Total revenue divided by quantity; (TR / Q)
Marginal revenue (MR)	Change in total revenue divided by change in quantity; (ΔTR / ΔQ)

We will study the various market structures in detail in the next reading. For the purposes of illustrating the different revenue terms we introduce perfect competition and imperfect competition at this point:

- In perfect competition, each individual firm faces a perfectly elastic demand curve (i.e., it can sell as many units of output as it desires at the given market price). The firm has no impact on market price, and is referred to as a price-taker.
- In imperfect competition, the firm has at least some control over the price at which it sells its output. The demand curve facing the firm is downward sloping so in order to increase units sold, the firm must lower its price. Stated differently, price and quantity demanded are inversely related. Firms operating in imperfect competition are referred to as price-searchers.

Total, Average, and Marginal Revenue Under Perfect Competition

Table 2-2 presents total, average, and marginal revenue for a firm that is a price-taker at each quantity of output (faces a perfectly elastic demand curve).

Table 2-2: TR, AR, and MR in Perfect Competition

Perfect Competition				
Qty.	**Price**	**TR**	**MR**	**AR**
0	8	0	0	–
1	8	8	8	8
2	8	16	8	8
3	8	24	8	8
4	8	32	8	8
5	8	40	8	8
6	8	48	8	8
7	8	56	8	8

Note: Under perfect competition, MR = Price
Therefore, the MR curve is the same as the demand curve.

Note that since the firm is a price-taker, price is fixed at $8 at all quantities of output sold.

2 - Exhibit 3, Volume 2, CFA Program Curriculum 2015

Prices are determined by demand and supply in the market. Once market price is determined, a firm in perfect competition can sell as many units of output as it desires at this price.

From Table 2-2 notice the following:

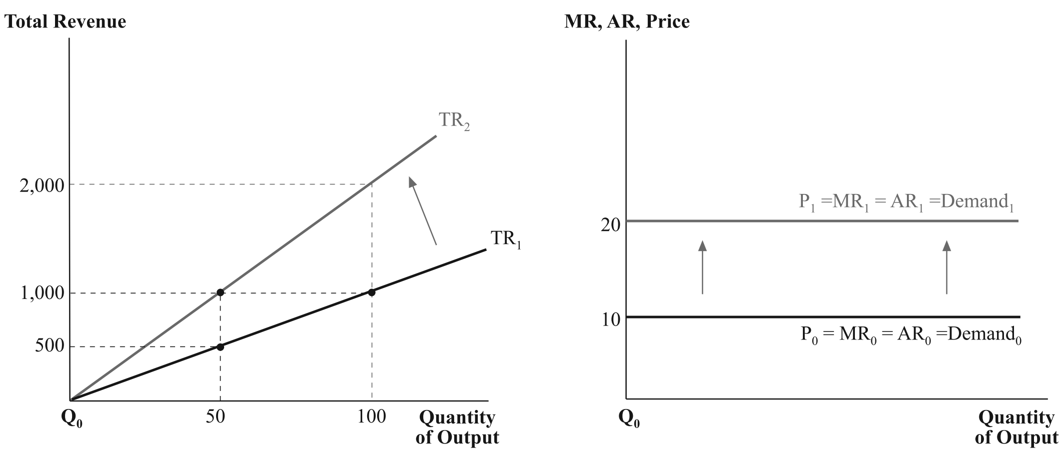
- Total revenue (TR) simply equals price times quantity sold. TR at 4 units is calculated as 4 × \$8 = \$32.
- Marginal revenue (MR) is defined as the increase in total revenue from selling one more unit. It is calculated as the change in total revenue divided by the change in quantity sold. MR from selling Unit 4 is calculated as (\$32 – \$24)/(4 – 3) = \$8.
- Average revenue (AR) equals total revenue divided by quantity sold. AR at 4 units of output is calculated as \$32/4 = \$8.

For any firm that sells all its output at a uniform price, average revenue will equal price regardless of the shape of the demand curve.

Important takeaways: In a perfectly competitive environment (where price is constant regardless of the quantity sold by the firm):

- MR always equals AR, and they both equal market price.
- If there is an increase in market demand, the market price increases, which results in both MR and AR shifting up (to MR_1 and AR_1 in Figure 2-1) and TR pivoting upward (to TR_1 in Figure 2-1).

Figure 2-1: TR, MR, and AR Under Perfect Competition



Total, Average, and Marginal Revenue Under Imperfect Competition

Table 2-3 presents total, average, and marginal revenue for a firm that is a price-searcher at each quantity of units sold (faces a downward-sloping demand curve).

Note that since the firm is a price-searcher, price and quantity are inversely related.

Table 2-3: TR, AR, and MR in Imperfect Competition

Price Searcher

Qty.	Price	TR	MR	AR
0	0	0	0	–
1	15	15	15	15
2	13	26	11	13
3	11	33	7	11
4	9	36	3	9
5	7	35	−1	7
6	5	30	−5	5
7	3	21	−9	3

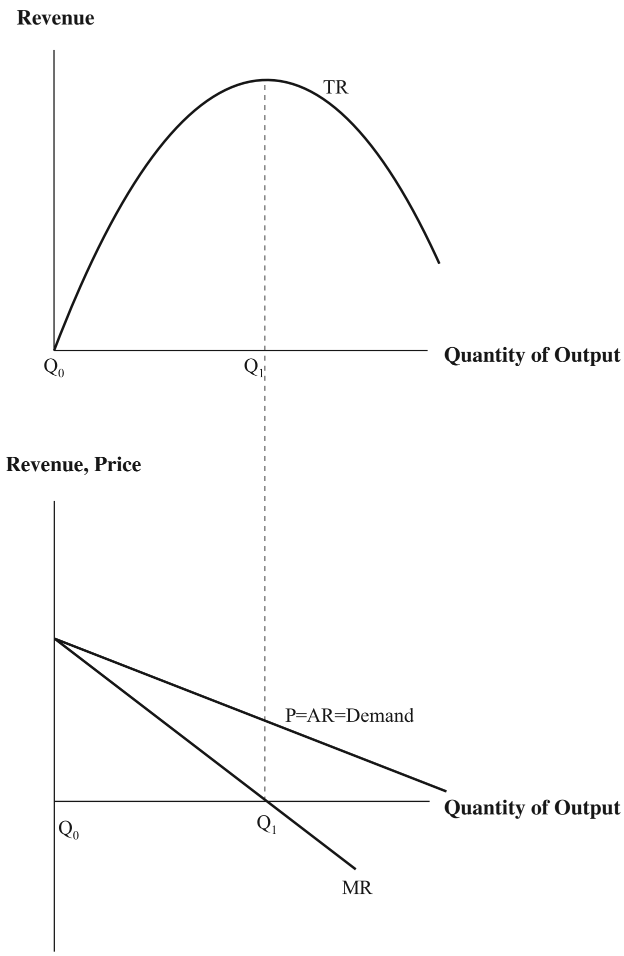
Note: Under imperfect competition, MR does not equal price. The MR curve has a steeper slope than the demand curve.

We mentioned earlier that for a firm that sells at a uniform price, average revenue will equal price. In Table 2-3 we have assumed that in order to increase quantity demanded and sold from 3 to 4 units, the firm must bring down its price from \$11 to \$9. The lower price (\$9) is applicable not only on the additional unit sold (the 4^{TH} unit) but also on all units that were previously selling for \$11. Only if the firm were a perfect monopolist would it be able to charge \$11 each for the first 3 units sold and \$9 for the 4^{TH} unit.

- Total revenue (TR) simply equals price times quantity sold. TR at 4 units is calculated as 4 × \$9 = \$36.
- Marginal revenue (MR) is defined as the increase in total revenue from selling one more unit. It is calculated as the change in total revenue divided by the change in quantity sold. MR from selling Unit 4 is calculated as (\$36 – \$33)/(4 – 3) = \$3. Under imperfect competition, in order to sell the 4^{TH} unit, the firm must entice further consumption by reducing its price to \$9. Moreover, not only does the buyer of the 4^{TH} unit pay a reduced price of \$9, but the 3 previous consumers also benefit from reduced prices and now only pay \$9 instead of \$11. On one hand, revenue increases by selling a larger quantity, (4^{TH} unit sells for \$9 resulting in an increase in TR of \$9). On the other hand, revenue falls (from selling the first three units at the new lower market price) by (\$11 – \$9) × 3 = \$6. The net increase in total revenue from selling the 4^{TH} unit equals \$9 – \$6 = \$3.
- Average revenue (AR) equals total revenue divided by quantity sold. AR at 4 units of output is calculated as \$36/4 = \$9.

Figure 2-2 illustrates TR, MR, and AR for a firm in imperfect competition.

Figure 2-2: TR, AR, and MR Under Imperfect Competition



Important takeaways: In imperfect competition (where price and quantity are inversely related):

- As quantity increases, the rate of increase in TR (as measured by MR) decreases.
- AR equals price at each output level.
- MR is also downward sloping with a slope that is steeper than that of AR (demand).
- TR reaches its maximum point when MR equals 0.

LOS 15c: Describe a firm's factors of production.

Factors of Production

The resources used by the firm in the production process are known as factors of production. They include:

- Land, the site location of the business.
- Labor, which includes skilled and unskilled labor, as well as managers.
- Capital (in this context physical capital), inputs used in the production process that are produced goods themselves (e.g., equipment and tools).
- Materials, goods purchased and used by the business as inputs to the production process.

In the following section, we assume that the firm only uses two factors of production, labor and capital. Therefore, the firm's production function is given as:

Q = f(K,L) where K ≥ 0 and L ≥ 0

Further, capital is assumed to be a fixed factor of production (the company cannot change the quantity of capital employed in its production process in the short run) while labor is a variable factor of production (the company can change the quantity of labor employed).

LOS 15d: Calculate and interpret total, average, marginal, fixed, and variable costs.

In this section we discuss the cost curve relationships in the short run (with labor as the only variable factor of production). In the long run all factors of production including technology, plant size, and physical capital are variable.

Total, Average, Marginal, Fixed, and Variable Costs

Table 2-4 summarizes various cost terms.

Table 2-4: Summary of Cost Terms[3]

Costs	Calculation
Total fixed cost (TFC)	Sum of all fixed expenses; here defined to include all opportunity costs
Total variable cost (TVC)	Sum of all variable expenses, or per unit variable cost times quantity; (per unit VC × Q)
Total costs (TC)	Total fixed cost plus total variable cost; (TFC + TVC)
Average fixed cost (AFC)	Total fixed cost divided by quantity; (TFC / Q)
Average variable cost (AVC)	Total variable cost divided by quantity; (TVC / Q)
Average total cost (ATC)	Total cost divided by quantity; (TC / Q) or (AFC + AVC)
Marginal cost (MC)	Change in total cost divided by change in quantity; (ΔTC / ΔQ)

Table 2-5 presents total costs, average costs, and marginal costs for a hypothetical company. We assume that the firm has rented 2 units of capital (Q_K) in the short run at $20 per unit. Labor is the only variable factor of production and the firm pays a wage of $10 per unit of labor employed. Even if the firm shuts down in the short run, it will still have to pay its total fixed costs (TFC), and if it wants to increase production in the short run, only total variable costs (TVC) will rise.

3 - Exhibit 13, Volume 2, CFA Program Curriculum 2015

Table 2-5: Total Cost, Average Cost, and Marginal Cost

Q_L	Q_K	TP	TFC	TVC	TC	AFC	AVC	ATC	MC
						$\frac{TFC}{TP}$		$\frac{TC}{TP}$	
0	2	0	40	0	40				
1	2	5	40	10	50	8	2	10	2
2	2	12	40	20	60	3.33	1.67	5	1.43
3	2	17	40	30	70	2.35	1.76	4.12	2
4	2	20	40	40	80	2	2	4	3.33
5	2	22	40	50	90	1.82	2.27	4.09	5
							$\frac{TVC}{TP}$		$\frac{\text{Change in TC}}{\text{Change in TP}}$

Total product (TP) is the maximum output a given quantity of labor can produce when working with a given quantity of capital units. Product terms (including total product) are discussed in detail later in the reading.

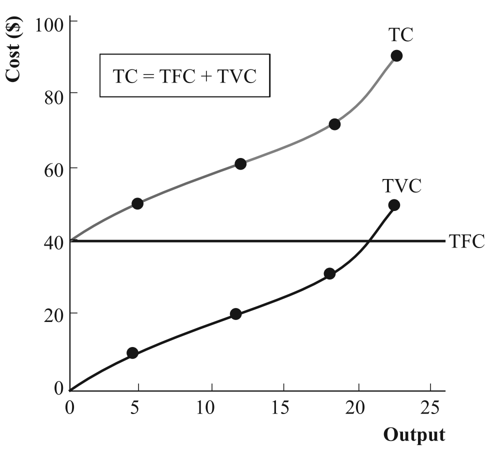
Total costs (TC) equal total fixed costs (TFC) plus total variable costs (TVC). Initially TC increases at a decreasing rate (green portion of TC curve in Figure 2-3). As production approaches full capacity TC increases at an increasing rate (grey portion of TC curve in Figure 2-3). At zero production, TC equals TFC as TVC equals 0.

Total fixed costs (TFC) equal the sum of all expenses that remain constant regardless of production levels. Since they cannot be arbitrarily reduced when production falls, fixed costs are the last expenses to be trimmed when a firm considers downsizing. Note that normal profit is also included in fixed cost.

Fixed costs include sunk costs. They also include quasi-fixed costs (e.g., utilities) that remain the same over a particular range of production, but move to another constant level outside of that range of production.

Total variable costs (TVC) are the sum of all variable costs. TVC is directly related to quantity produced (TP), and the shape of the TVC curve mirrors that of the TC curve. Whenever a firm looks to downsize or cut costs, its variable costs are the first to be considered for reduction as they vary directly with output.

Figure 2-3: Cost Curves



Notice in Figure 2-3 that the TC and TVC increase at a decreasing rate at low levels of output, and increase at an increasing rate at higher levels of output. The difference between TC and TVC equals TFC.

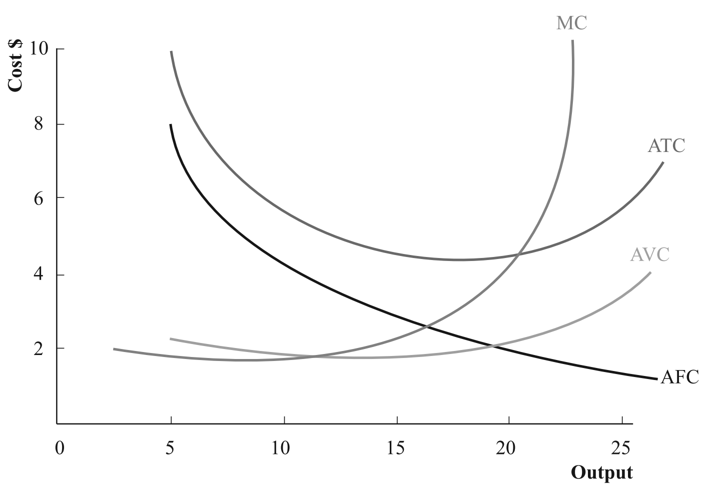
Average total cost (ATC) is simply total cost (TC) divided by total product (TP).

Average fixed cost (AFC) equals TFC divided by TP.

Average variable cost (AVC) equals TVC divided by TP.

Marginal cost (MC) equals the increase in total costs brought about by the production of one more unit of output. It equals change in total costs divided by the change in output (TP). Since TFC is fixed, MC can also be calculated as the change in TVC divided by the change in TP. Though not clearly visible from Figure 2-4, we should emphasize that the MC curve is shaped like a "J."

Figure 2-4: Marginal and Average Cost Curves



A firm's MC is the increase in total cost from producing the last unit of output. For example, when output increases from 12 to 17 units (Table 2-5), TC increases by $10, to $70. The MC of any of these 5 units (the 13TH to 17TH units) equals $10 divided by 5, or $2.

It is important to bear in mind that MC illustrates the slope of the TC curve at a particular level of output. MC initially decreases (see Figure 2-4) because of the benefits from *specialization*. However, MC eventually increases because of *diminishing marginal returns*. To produce more output given the same amount of capital, more and more units of labor must be employed because each additional unit of labor is less productive than the previous one. Since more workers are required to produce one more unit of output, the cost of producing that additional unit (the marginal cost) increases.

From Figure 2-4, also notice that:

- As output levels rise, total fixed costs are spread over more and more units and AFC continues to fall at a decreasing rate.
- The average total cost curve is U-shaped. It falls initially as fixed costs are spread over an increasing number of units. Later however, the effect of falling AFC is offset by diminishing marginal returns so ATC starts rising.

- The vertical distance between the AVC and ATC curves equals AFC. This vertical distance between the AVC and ATC curves gets smaller as output increases because AFC decreases as output expands.
- The minimum point of the AVC does not correspond to the minimum point of ATC. The firm's profit-maximizing quantity does not necessarily occur at the point where ATC is minimized, even though profit per unit may be maximized at this point.

Important Relationships between Average and Marginal Cost Curves

- MC intersects ATC and AVC from *below* at their respective *minimum* points.
- When MC is below AVC, AVC falls, and when MC is above AVC, AVC rises.
- When MC is below ATC, ATC falls, and when MC is above ATC, ATC rises.

LOS 15e: Determine and describe breakeven and shutdown points of production.

The Firm's Short Run Supply Curve

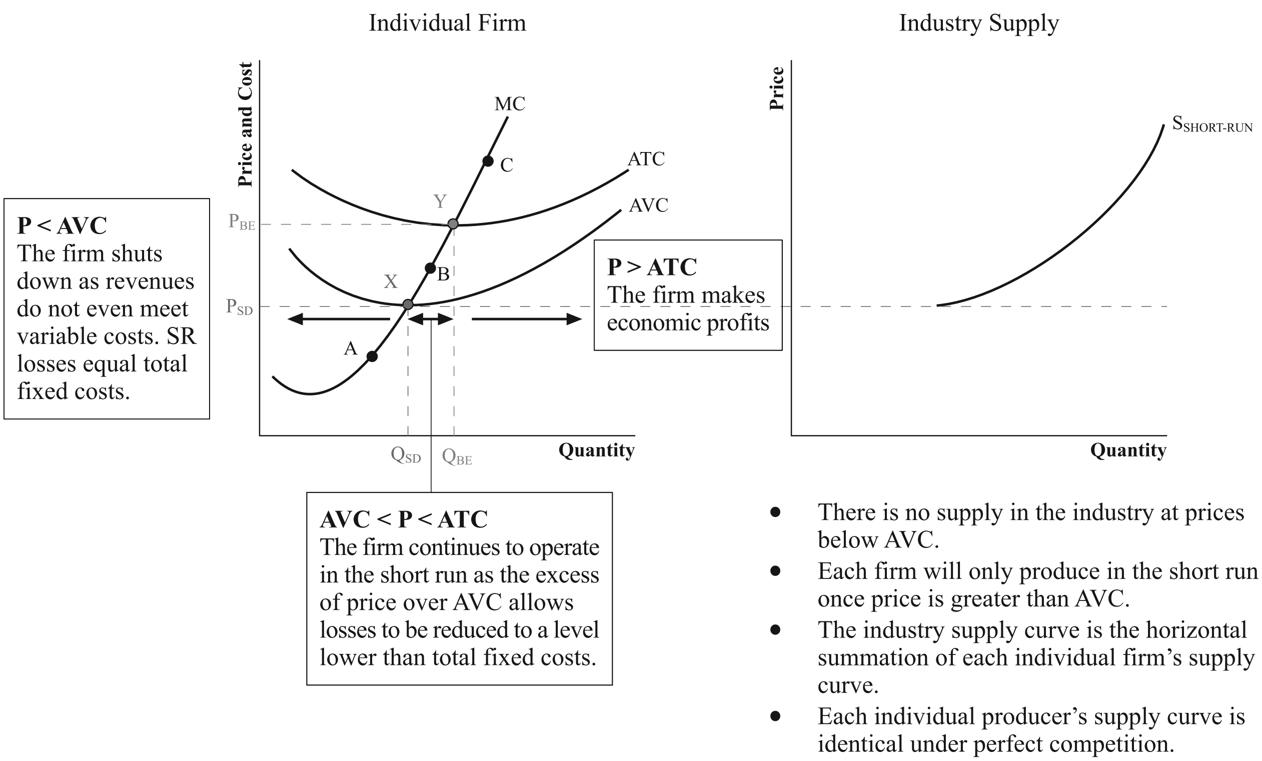
In the short run, a firm incurs fixed and variable costs of production. If the firm decides to shut down, it will still incur fixed costs in the short run and make a loss equal to total fixed cost. This loss can be reduced by continuing production and earning revenues that are greater than the variable costs of production. This surplus (excess of revenues over variable costs) would serve to meet some of the fixed costs that the firm is "stuck" with in the short run. A firm should shut down immediately if it does not expect revenues to exceed variable costs of production. If the firm continues to operate in such an environment, it would suffer a loss greater than just total fixed cost.

To make the analysis easier to understand, you might want to think about price as average revenue. For example, at Point A, average revenue would be less than average variable cost. Therefore, total revenue would be less than total variable costs and the firm would be better off shutting down and limiting losses to total fixed costs.

Therefore, an individual firm's short run supply curve (that illustrates its willingness and ability to produce at different prices) is the portion of its MC curve that lies above the AVC curve (see Figure 2-5).

- At price levels below AVC (e.g., Point A where P < AVC), the firm will not be willing to produce, as continued production would only extend losses beyond simply total fixed costs. Any quantity to the left of Point X (with quantity Q_{SD}) would define a shutdown point for the firm.
- When price lies between AVC and ATC (e.g., Point B where AVC < P < ATC), the firm will remain in production in the short run as it meets all variable costs and covers a portion of its fixed costs.
- To remain in business in the long run, the firm must breakeven or cover all costs (revenues should meet total costs). Therefore, Point Y defines the firm's breakeven point. In the long run, at any price lower than P_{BE}, the firm will exit the industry.
- Once prices exceed ATC (e.g., Point C where P > ATC) the firm makes economic profits.

Figure 2-5: SR Supply, Shutdown, and Breakeven Under Perfect Competition

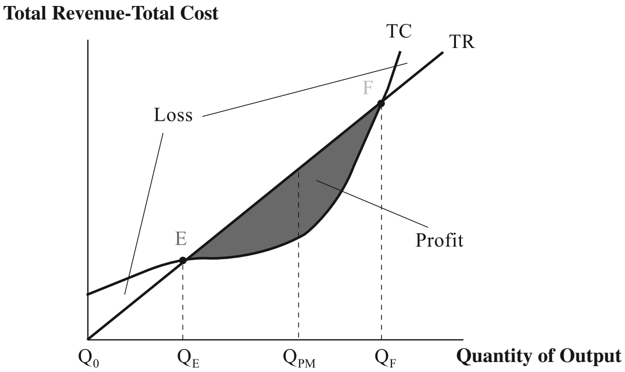


Note: In perfect competition, the breakeven point will occur at the quantity where MR = AR = P_{BE}. Further, all three will equal minimum ATC.

We can also define the firm's breakeven point under perfect competition using the total revenue-total cost approach. Figure 2-6 presents two points where TC and TR are equal (intersection points).

- Below the lower intersection point, Point E, the firm is making a loss as TC is greater than TR.
- Above the upper intersection point, Point F, the firm is making a loss as TC is greater than TR.
- In between these two points, the firm makes economic profits, with profits being maximized where the vertical distance between TR and TC is maximized. The firm would not want to expand production beyond Q_{PM} as this is where it maximizes its profit.
- The higher the initial breakeven point, the more risky the business as it would take a higher volume to reach initial breakeven. However, once the business starts making profits (at higher output levels) it should expect to attain higher returns to compensate for the higher risk.

Figure 2-6: Profit Maximization Using the Total Revenue-Total Cost Approach



If the firm's TC curve is always higher than its TR curve, the firm will want to minimize its economic losses. This occurs at the point where the vertical distance between TC and TR is minimized (Point G corresponding to Q_{LM} in Figure 2-7).

Figure 2-7: Loss Minimization Using the Total Revenue-Total Cost Approach

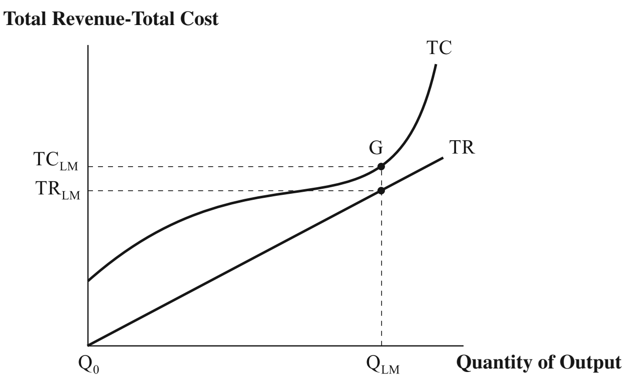


Table 2-6 summarizes the above analysis.

Table 2-6: Operating Decisions[4]

Revenue-Cost Relationship	Short-Run Decision	Long-Term Decision
TR ≥ TC	Stay in market	Stay in market
TR > TVC but TR < TFC + TVC	Stay in market	Exit market
TR < TVC	Shut down production to zero	Exit market

4 - Exhibit 19, Volume 2, CFA Program Curriculum 2015

LESSON 3: MAXIMIZING PROFITS AND OPTIMIZING PRODUCTIVITY

LOS 15f: Describe approaches to determining the profit-maximizing level of output.

There are three approaches to determining the output level at which profits are maximized.

1. The point where the difference between total revenue and total costs is maximized.
2. The point where the last unit sold adds as much to revenue as it does to costs (i.e., the last unit of output breaks even). Stated differently, profits are maximized at the output level where MC equals MR. If MR exceeds MC, profits rise if production is increased. If MR is less than MC, profits fall (rise) if production is increased (decreased).
3. The point where the revenue from the last input unit equals the cost of that unit. If the increase in revenue from the employment of an additional input unit exceeds its cost, profits rise from employment of that unit. If the increase in revenue from the employment of the additional input unit is lower than its cost, profits would fall from employment of that unit.

Profit Maximization Under Perfect Competition

Table 3-1 presents cost and revenue information for a firm in perfect competition.

Table 3-1: Profit Maximization Under Perfect Competition

Q	P	TR	MR	MC	FC	TC	Profit = TR – TC
0	8	0	0	0	15	15	–15
1	8	8	8	0.50	15	15.50	–7.50
2	8	16	8	1.25	15	16.75	–0.75
3	8	24	8	2.15	15	18.90	5.10
4	8	32	8	4	15	22.90	9.10
5	8	40	8	6	15	28.90	11.10
6	8	48	8	9	15	37.90	10.10
7	8	56	8	13	15	50.90	5.10
8	8	64	8	15	15	65.90	–1.90

From Table 3-1, notice the following:

- At zero production, the firm incurs an economic loss of $15 (which equals TFC).
- The firm breaks even by Unit 3.
- The region of profitability ranges from 3 units to 7 units.
- Total profit is maximized at 5 units, where profits amount to $11.10.
- At this 5-unit output level, MR exceeds MC ($8 versus $6).When production expands to 6 units, MR falls below MC ($8 versus $9). The 6^{TH} unit makes a loss of $1, so profit drops from $11.10 to $10.10.
- Beyond 7 units, the firm enters the second economic loss zone.

Profit Maximization Under Imperfect Competition

Table 3-2 presents cost and revenue information for a firm in imperfect competition.

Since price does not equal marginal revenue or average revenue for a firm in imperfect competition, analysis of shut down and breakeven points should be based on total costs and revenues.

Table 3-2: Profit Maximization Under Imperfect Competition

Q	P	TR	MR	MC	FC	TC	Profit = TR – TC
0	11	0	0	0	15	15	–15
1	10	10	10	0.50	15	15.50	–5.50
2	9	18	8	1.25	15	16.75	1.25
3	8	24	6	2.15	15	18.90	5.10
4	7	28	4	4	15	22.90	5.10
5	6	30	2	6	15	28.90	1.10
6	5	30	0	9	15	37.90	–7.90
7	4	28	–2	13	15	50.90	–22.90
8	3	24	–4	15	15	65.90	–41.90

From Table 3-2, notice the following:

- At zero production, the firm incurs an economic loss of $15 (which equals TFC).
- The firm breaks even by Unit 2.
- The region of profitability ranges from 2 units to 5 units.
- Total profit is maximized at 4 units, where profit equals $5.10.
- At this 4-unit output level, MR equals MC ($4).When production expands to 5 units, MR falls below MC ($2 versus $6). The 5TH unit makes a loss of $4 so profit drops from $5.10 to $1.10.
- Beyond 5 units, the firm enters the second economic loss zone.

Table 3-3 summarizes the TR-TC and MR-MC profit-maximization approaches for a firm under perfect competition.

Table 3-3: Profit Maximization Under Perfect Competition[5]

Revenue–Cost Relationship	Actions by Firm
TR = TC and MR > MC	Firm is operating at lower breakeven point; increase Q to enter profit territory.
TR ≥ TC and MR = MC	Firm is at maximum profit level; no change in Q.
TR < TC and TR ≥ TVC but (TR – TVC) < TFC (covering TVC but not TFC)	Find level of Q that minimizes loss in the short run; work toward finding a profitable Q in the long run; exit market if losses continue in the long run.
TR < TVC (not covering TVC in full)	Shut down in the short run; exit market in the long run.
TR = TC and MR < MC	Firm is operating at upper breakeven point; decrease Q to enter profit territory.

5 - Exhibit 26, Volume 2, CFA Program Curriculum 2015

LOS 15j: Calculate and interpret total, marginal, and average product of labor.

LOS 15k: Describe the phenomenon of diminishing marginal returns and calculate and interpret the profit-maximizing utilization level of an input.

LOS 15l: Determine the optimal combination of resources that minimizes cost.

Productivity

In the short run, at least one factor of production is fixed. Usually we assume that labor is the only variable factor of production in the short run. Therefore, the only way that a firm can respond to changing market conditions in the short run is by changing the quantity of labor that it employs. In hard times, firms lay off labor and in good times, firms employ more labor. However, the quantities of capital and land employed remain fixed. A firm cannot increase output in the short run by acquiring more machinery or equipment.

In the long run, quantities of all factors of production can be varied. Output can be increased by employing more labor, acquiring more machinery or even buying a whole new plant. However, once a long-run decision has been made (e.g., the acquisition of a new plant), it cannot be easily reversed.

Total, Average, and Marginal Product of Labor

In Table 3-4 we assume that in the short run, the firm invests in 2 units of capital, which comprise its fixed costs. The only factor of production whose quantities it can vary is labor. As more units of labor are employed to work with 2 units of capital, total output increases.

Table 3-4: Total Product, Marginal Product, and Average Product

			$\frac{\text{Change in TP}}{\text{Change in } Q_L}$	$\frac{TP}{Q_L}$
Quantity of Labor (Q_L)	**Quantity of Capital (Q_K)**	**Total Product (TP)**	**Marginal Product (MP)**	**Average Product (AP)**
0	2	0		
1	2	5	5	5.00
2	2	12	7	6.00
3	2	17	5	5.67
4	2	20	3	5.00
5	2	22	2	4.4
6	2	21	−1	3.5

Total product (TP) is the maximum output that a given quantity of labor can produce when working with a fixed quantity of capital units. While TP provides an insight into the company's size relative to the overall industry, it does not show how efficient the firm is in producing its output.

Marginal product (MP) (also known as marginal return) equals the increase in total product brought about by hiring one more unit of labor, while holding quantities of all other factors of production constant. MP measures the productivity of the individual additional unit of labor.

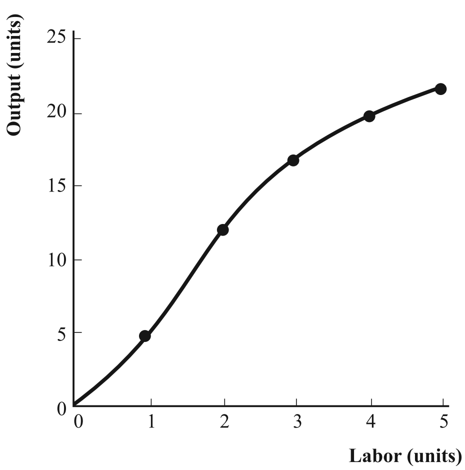
Average product (AP) equals total product of labor divided by the quantity of labor units employed. AP is a measure of overall labor productivity. The higher a firm's AP, the more efficient it is.

AP and MP provide valuable insights into labor productivity. However, when individual worker productivity is difficult to monitor, (e.g., when tasks are performed collectively) AP is the preferred measure of labor productivity.

From Table 3-4, notice that TP continues to increase until the 6^{TH} unit of labor is employed. The firm would obviously not hire a unit with negative productivity so only the first 5 units of labor are considered for employment.

Figure 3-1 illustrates the firm's total product (TP) curve. In the initial stages (as the first and second units of labor are employed), total product *increases at an increasing rate*. The slope of the total product curve is relatively *steep* at this stage. Later, as more units of labor are employed to work with the fixed 2 units of capital, total output *increases at a decreasing rate*, and the slope of the TP curve becomes *flatter*.

Figure 3-1: Total Product



The marginal product (MP) curve (see Figure 3-2) shows the change in total product from hiring one additional unit of labor. The MP curve is simply the slope of the TP curve. Recall that MP is calculated as the change in TP (change on y-axis) divided by the change in labor units (change on x-axis). The MP curve rises initially, (over the first two units of labor when TP is increasing at an increasing rate) and then falls (as the 3^{RD}, 4^{TH}, and 5^{TH} units of labor are employed and TP increases at a decreasing rate). If the 6^{TH} unit is employed, MP turns negative and TP falls.

Diminishing marginal returns may also set in as the most productive units of labor are hired initially, and then as the firm seeks to increase production, less competent units of labor must be employed.

The firm benefits from increasing marginal returns over the first two units of labor (as MP increases) and then suffers from decreasing (or diminishing) marginal returns over the next four units of labor (as MP decreases). Increasing marginal returns occur because of specialization and division of labor, while decreasing marginal returns set in because of inefficiency, over-crowdedness, and underemployment of some units of labor given the fixed amount of capital.

The average product (AP) curve (see Figure 3-2) shows output per worker, which equals total product divided by total quantity of labor. Observe two important relationships from the AP and MP curves:

1. MP intersects AP from *above* through the *maximum point* of AP.
2. When MP is *above* AP, AP *rises*, and when MP is *below* AP, AP falls.

An interesting way to remember the relationship stated in Point 2 is by analyzing the historical returns earned by a fund manager. If her 5-year average return is 10%, and she earns 15% this year (marginal return), her average would rise. If however, she were to earn only 5% this year, her average would fall.

Figure 3-2: Average and Marginal Product Curves

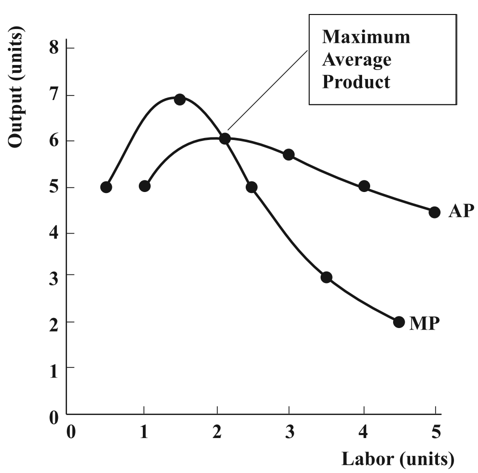
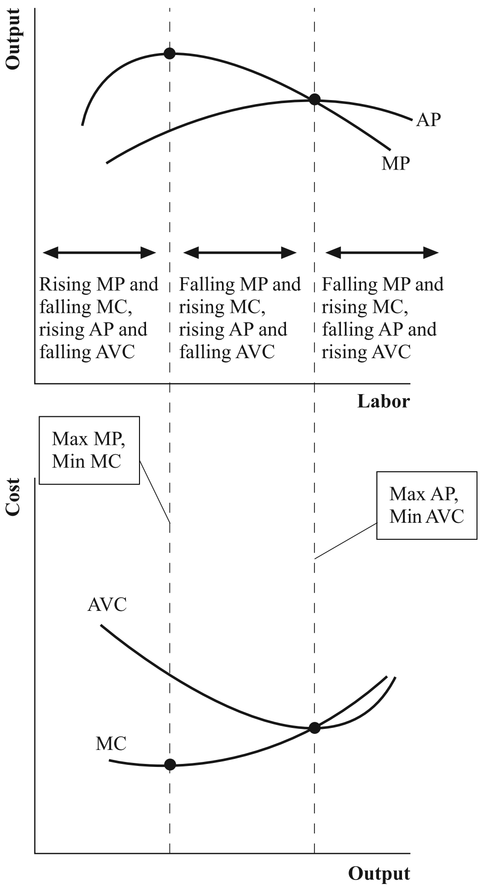


Figure 3-3 illustrates important relationships between cost and product curves.

Figure 3-3: Cost and Product Curves



- A firm's MP curve is linked to its MC curve. Initially, as more labor is hired, MP rises and MC falls.
- At the point where MP reaches its maximum, MC stands at its minimum.
- As output expands further, MP falls and MC rises.

- A firm's AP curve is linked to its AVC curve. Initially, as the firm hires more labor, AP rises and AVC falls.
- At the point where AP reaches its maximum, AVC is at its minimum.
- As the firm increases output further, AP falls and AVC rises.

The productivity of different input factors is compared on the basis of output per unit of input cost (MP_{input}/P_{input}). If a firm uses a combination of labor and capital, the least-cost optimization formula would be given by the following equation:

$$\boxed{\frac{MP_L}{P_L} = \frac{MP_K}{P_K}}$$

If MP_L/P_L equals 3 and MP_K/P_K equals 6, it implies that physical capital offers two times the output per money unit of input cost relative to labor. Given these ratios, if the firm plans to increase production, it would choose to employ physical capital over labor as every additional dollar spent on capital yields two times the output of an additional dollar spent on labor. As the firm adds more units of physical capital, the MP of capital declines as diminishing marginal returns to capital set in. This would decrease the value of MP_K/P_K. The firm would keep adding units of capital until MP_L/P_L and MP_K/P_K are equal (both equal 3).

This analysis assumes that MP_L remains constant as Q_K changes. Practically speaking, we would expect MP_L to rise as Q_K increases because labor would become more productive using more physical capital. Eventually the values of MP_K/P_K and MP_L/P_L would actually end up somewhere between 3 and 6.

In order to determine the profit-maximizing quantity of an input unit, we compare the revenue value of the unit's MP to the cost of the unit. Marginal revenue product (MRP) of labor measures the increase in total revenue from selling the additional output (marginal product) produced by the last unit of labor employed. It is calculated as follows:

MRP of labor = Change in total revenue / Change in quantity of labor

For a firm in perfect competition, MRP of labor equals the MP of the last unit of labor times the price of the output unit.

MRP = Marginal product × Product price

A profit-maximizing firm will hire more labor until:

$$\text{MRP}_{\text{Labor}} = \text{Price}_{\text{Labor}}$$

- If the price of labor, or wage rate, is *less* than the marginal revenue product of labor, a firm can increase its profit by employing one *more* unit of labor.
- If the wage rate is *greater* than the marginal revenue product of labor, profit can be increased by employing one *less* unit of labor.

In case the firm uses more than one factor of production, profits are maximized when the MRP of each factor equals the price of each factor unit. This means that the ratio of the factor's MRP to its price should equal 1.

Profits are maximized when:

$$\boxed{\frac{\text{MRP}_1}{\text{Price of input 1}} = \ldots = \frac{\text{MRP}_n}{\text{Price of input n}} = 1}$$

Example 3-1: Profit Maximization Using the MRP-Factor Cost Approach

The table below lists MRP, compensation, and $\text{MRP}_{\text{input}}/\text{P}_{\text{input}}$ information for skilled, semi-skilled, and unskilled labor.

Type of Labor	$\text{MRP}_{\text{input}}$/Day	$\text{Compensation}_{\text{input}}$/Day	$\text{MRP}_{\text{input}}/\text{P}_{\text{input}}$
Unskilled (U)	50	50	1.0
Semi-skilled (SS)	200	100	2.0
Skilled (S)	360	120	3.0

Given that the quality of output produced by all three types of labor units is exactly the same, determine which type of labor contributes most to profits.

Solution:

Since the MRP_{input}/P_{input} ratio is highest for skilled labor, the firm adds skilled labor first as it is the most profitable type of labor to employ. The contribution made by the additional unit of skilled labor to total profit is calculated as the difference between its MRP and price ($360 – $120 = $240). However, as more units of skilled labor are employed, diminishing marginal returns set in (MP of skilled labor falls) so the MRP of skilled labor declines. Once the MRP_{input}/P_{input} ratio for skilled labor falls below 2.0, it becomes more feasible for the firm to hire additional units of semi-skilled labor rather than skilled labor as the contribution to profits of semi-skilled labor now exceeds that of skilled labor. Note that the firm would not hire any additional units of unskilled labor as MRP_U/P_U is already 1 and the next unit of unskilled labor would probably reduce profit (as its MP would fall, dragging the MRP_U/P_U ratio below 1.0). The input level that maximizes profit is where:

$MRP_U/P_U = MRP_{SS}/P_{SS} = MRP_S/P_S = 1$

LESSON 4: ECONOMIES AND DISECONOMIES OF SCALE AND PROFIT MAXIMIZATION IN THE SHORT RUN V. LONG RUN

LOS 15g: Describe how economies of scale and diseconomies of scale affect costs.

The Production Function

We defined the short run and long run in the previous section. Stated briefly, at least one factor of production is fixed in the short run, while no factors of production are fixed in the long run. Table 4-1 illustrates a company's production function, which shows how different quantities of labor and capital affect total product. In our short-run scenario, different quantities of labor were combined with fixed quantities of capital (2 units). Now we explore the effects of varying capital quantities (Q_K) as well (note: no factor of production is fixed, so we are basically working with the long-run scenario here). Plant 1 has 2 units of capital (the SR scenario that we analyzed earlier), Plant 2 has 4 units of capital, Plant 3 has 6 units of capital, and Plant 4 has 8 units of capital.

Table 4-1: A Firm's Production Function

Units of Labor	Plant 1 $Q_K = 2$	Plant 2 $Q_K = 4$	Plant 3 $Q_K = 6$	Plant 4 $Q_K = 8$
1	5	11	15	18
2	12	25	34	39
3	17	36	48	56
4	20	42	57	67
5	22	45	59	68

Diminishing Marginal Returns to Capital

Diminishing Marginal Returns to Labor

The effect of diminishing marginal returns to labor is obvious. Marginal product declines as more and more units of labor are added to each plant (move down the columns in Table 4-1). But now, we also see diminishing marginal returns to capital (see Table 4-2, and move right along the rows in Table 4-1). For example, with 3 units of labor, moving from Plant 1 to Plant 2 (increasing quantity of capital by 2 units) increases production by 19 units from 17 to 36 (MP of capital = change in output/change in quantity of capital = 19/2 = 9.5). Adding a further 2 units of capital and moving to Plant 3 increases production by only 12 units from 36 to 48 (MP of capital = change in output/change in quantity of capital = 12/2 = 6). Finally, adding another 2 units of capital and moving to Plant 4 and holding quantity of labor constant at 3 units increases output only by 8 units (MP of capital = 4).

Table 4-2: Diminishing Marginal Returns to Capital

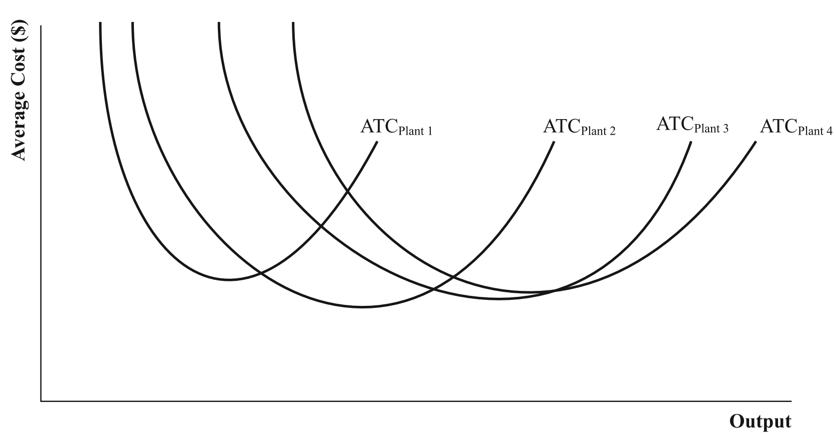
	Units of Capital Combined with 3 Units of Labor	Total Product when $Q_L = 3$	Marginal Product
Plant 1	2	17	8.5
Plant 2	4	36	9.5
Plant 3	6	48	6
Plant 4	8	56	4

Diminishing Marginal Returns to Capital

When we map the average cost curves for all 4 plants on one graph (see Figure 4-1), notice that:

- Short-run ATC curves are U-shaped.
- The larger the plant, the greater the output at which short-run ATC is at its minimum.

Figure 4-1: Average Cost Curves for Different Plant Sizes



The optimal output level for each plant is when its ATC curve is at its minimum. The long-run average cost (LRAC) curve illustrates the relationship between the lowest attainable average total cost and output when all factors of production are variable. The LRAC curve is also known as a planning curve because it shows the expected per-unit cost of producing various levels of output using different combinations of factors of production.

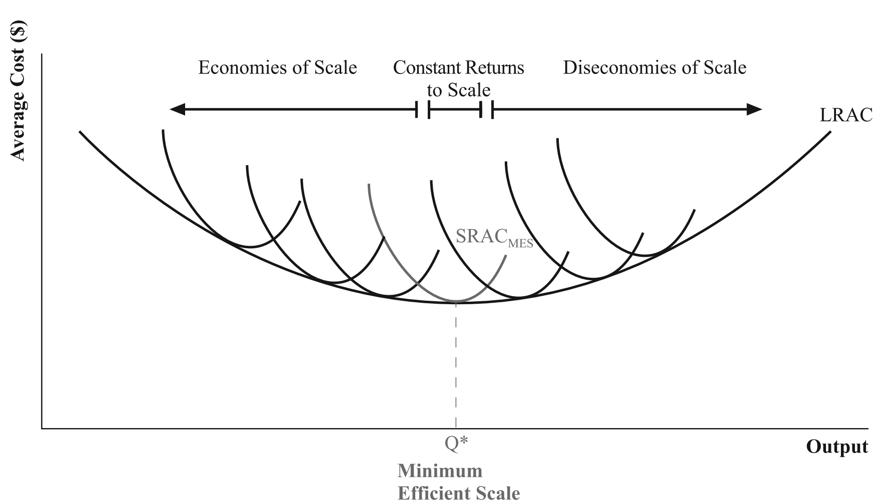
Economies and Diseconomies of Scale

Economies and diseconomies of scale are long-run concepts. They relate to conditions of production when all factors of production are variable. In contrast, increasing and diminishing marginal returns are short-run concepts, applicable only when the firm has one variable factor of production.

Economies of scale or increasing returns to scale refer to reductions in the firm's average costs that are associated with the use of larger plant sizes to produce large quantities of output. They are present over the range of output when the LRAC curve is falling. Economies of scale occur because mass production is more economical, the specialization of labor and equipment improves productivity, and costs such as advertising can be spread across more units of output. Other reasons include discounted prices as a result of bulk purchasing of resources and the ability to adopt more expensive but more efficient technology. When a firm is operating in the economies of scale region of the LRAC curve (see Figure 4-2), it should aim to expand capacity to enhance competitiveness and efficiency.

Diseconomies of scale or decreasing returns to scale occur in the upward-sloping region of the LRAC curve. A typical reason for an increase in average costs as output levels rise is an increase in bureaucratic inefficiencies as effective management, supervision, and communication become difficult in large organizations. Other reasons include duplication of business functions and product lines, and high resource prices due to supply constraints. When a firm is operating in the diseconomies of scale region of the LRAC curve (see Figure 4-2), it should aim to downsize and reduce costs to increase competitiveness.

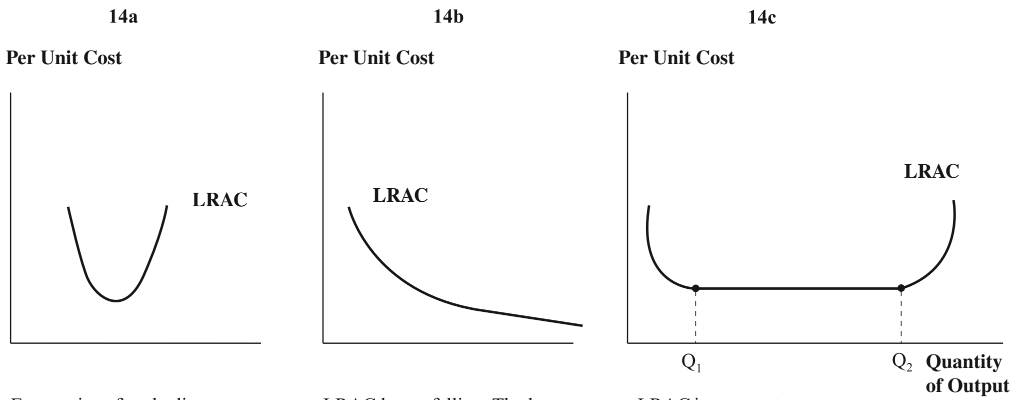
Figure 4-2: The Planning Curve



In the horizontal portion of the LRAC curve, when an increase in output does not result in any change in average costs, a firm realizes constant returns to scale.

Aside from the shape illustrated in Figure 4-2, the LRAC curve may also take one of the shapes described in Figure 4-3.

Figure 4-3: Types of LRAC Curves



Economies of scale disappear rapidly. A firm producing a small volume can actually be more efficient than one that produces a higher volume.

LRAC keeps falling. The larger the business the more efficient it becomes.

LRAC is constant over a range of output. Size does not give the firm a competitive edge on its rivals in this range.

Under perfect competition, given the different SRAC options available to the firm, $SRAC_{MES}$ embodies the optimal combination of technology, plant capacity, capital, and labor that minimizes the firm's average costs in the long run. The lowest point on the LRAC curve is called the firm's minimum efficient scale. In the long run, all firms in perfect competition operate at their minimum efficient scale as price equals minimum average cost (see next section).

LOS 15h: Distinguish between short-run and long-run profit maximization.

Equilibrium in the Short Run

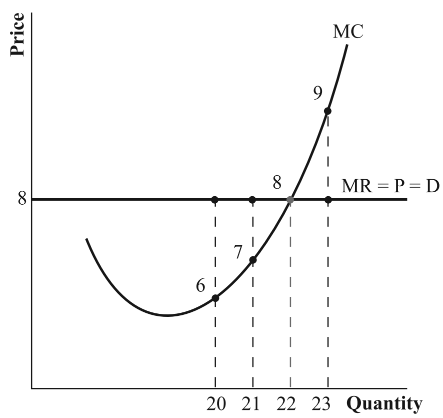
In perfect competition, each firm "takes" the price offered by the market, so the only decision in each producer's hands is how much to produce. In Figure 4-4 we illustrate an individual firm's marginal cost curve and demand curve.

In a perfectly competitive environment, an individual firm can sell as many units of output as it desires at the given market price. If the price is $8/unit, each unit sold increases revenue by $8, so MR is constant at the same level as price ($8).

The key decision for the firm is to decide on how much it wants to produce in the short run. This decision is governed by the incentive to maximize profits. In Figure 4-4, profit-maximizing output occurs at 22 units where the increase in revenue from selling the last unit (MR) equals the increase in costs from producing the last unit (MC).

Profit maximizing output is the quantity where the difference between TR and TC is maximized. This level of output corresponds to the point where MC equals MR. At production levels lower than 22 units, an additional unit produced adds more to revenues than to costs. For example, the 20^{TH} unit adds \$8 to revenue, \$6 to costs, and therefore \$2 to total profit. At production levels greater than 22 units, the cost of producing an extra unit exceeds the revenue from selling it. For example, the 23^{RD} unit costs \$9 and sells for only \$8. There is no point in increasing production beyond 22 units given current prices and the firm's cost structure.

Figure 4-4: Profit-Maximizing Output



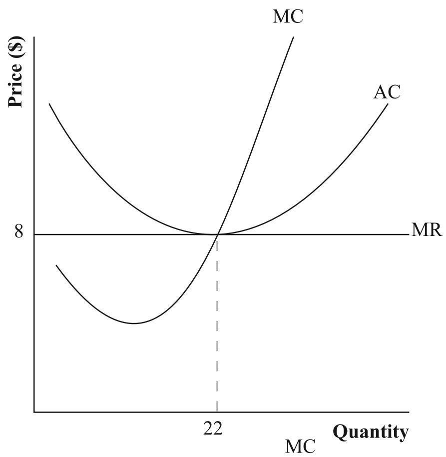
MR equals price only for price takers.

The 20^{TH} unit increases revenue by \$8 (selling price) and costs \$6 (marginal cost) to make. The increase in profit from 20^{TH} unit = \$2.

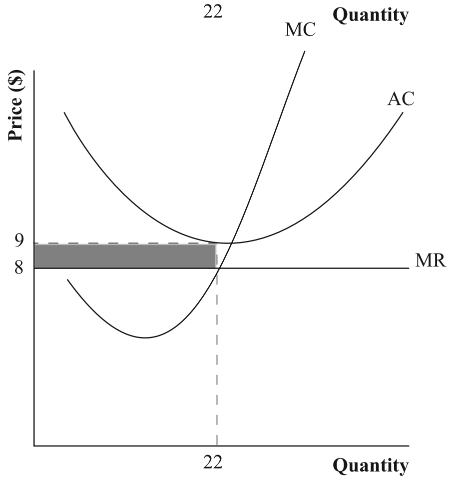
Unit	**Price/MR**	**MC**	**Profit on Unit**
20^{TH}	8	6	2
21^{ST}	8	7	1
22^{ND}	8	8	0
23^{RD}	8	9	–1

Whether the firm makes a profit or a loss depends on the position of the AC curve relative to demand (which represents average revenue). Three possible scenarios are illustrated in Figure 4-5. If the average cost curve lies above the firm's demand curve, the firm will make economic losses (see Figure 4-5b). However, the point where marginal cost equals marginal revenue will continue to be the quantity that the firm should produce, but in this case, the corresponding quantity will define the *loss-minimizing* level of output.

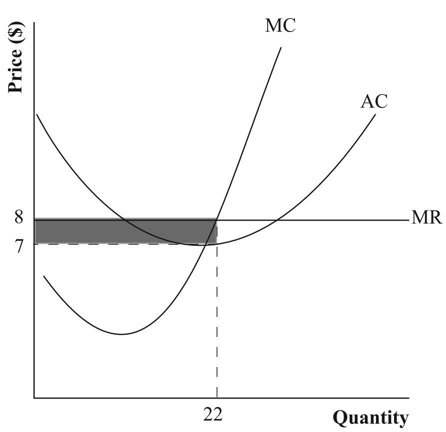
Figure 4-5: Possible Perfect Competition Scenarios in the Short Run



4-5a
Scenario A: Short Run Breakeven ⇨ Zero Economic Profit
MC equals MR at the point where AC is at its minimum and the AC curve is tangent to the demand curve. Firms only make normal profits as price equals AC.



4-5b
Scenario B: Short Run Economic Losses
At the profit maximizing quantity where MC = MR (22 units) price is lower than average cost (AC). Therefore, the firm makes an economic loss. The firm earns \$8/unit and spends \$9/unit so loss/unit equals \$1.
Total loss equals (\$8 – \$9) × 22 = \$22, which is the area of the rectangle shaded in grey.



4-5c
Scenario C: Short-Run Economic Profits
At the quantity where MC equals MR (22 units), the firm's price/unit is greater than its cost/unit (AC) so it makes an economic profit. The firm still produces 22 units. It earns \$8/unit and spends \$7/unit so profit/unit equals \$1. Total profit equals (\$8 – \$7) × 22 = \$22, which is the area of the rectangle shaded in green.

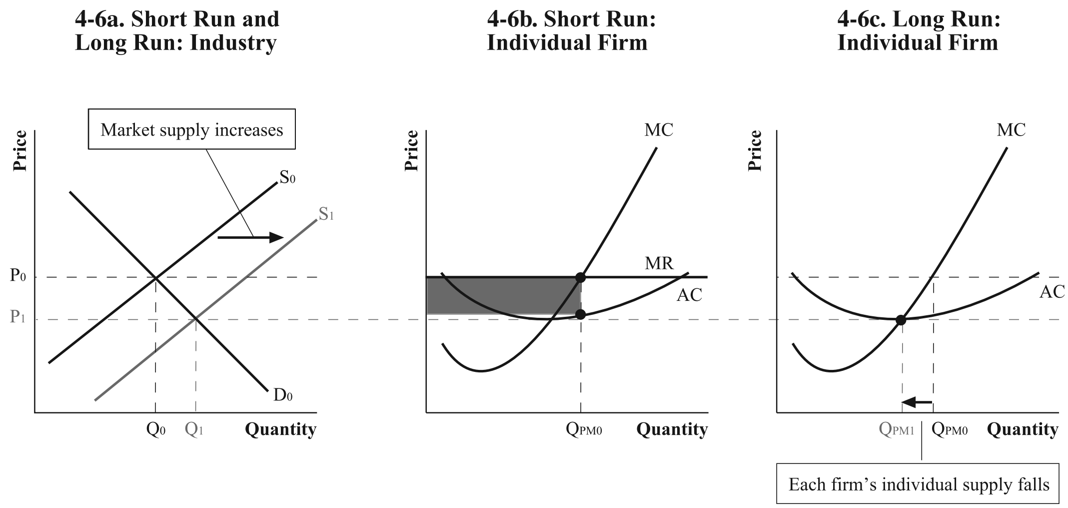
Equilibrium in the Long Run

The three scenarios illustrated in Figure 4-5 are only possible in the short run. In the long run, in a perfectly competitive industry, only one of those scenarios is possible (zero economic profits). In perfect competition (as you will see in the next reading) there are no barriers to entry or exit. Firms can enter the industry easily when they see persistent economic profits, and may leave if they foresee persistent economic losses. Let's see what happens in each scenario.

Economic Profits

If firms in a particular industry are making economic profits, entrepreneurs will flock to the industry to capture some of the economic profits available, as shown in Figure 4-6. Note that established firms enjoy no advantage over new entrants under perfect competition and each firm has an identical cost structure. The increase in the number of firms increases market supply to S_1 (see Figure 4-6a). Consequently, market prices fall till they reach the point where price equals minimum ATC and economic profits are eliminated (P_1).

Figure 4-6: SR Economic Profits, LR Zero Economic Profits



SR: S_0 and D_0 interact to set prices at P_0.

LR: Other firms are drawn to the industry so industry supply gradually moves to S_1 to set LR market prices at P_1.

SR: Each firm produces Q_{PM0} where MR = MC.

Prices exceed AC at Q_{PM0} resulting in economic profits (region shaded in green) equal to $(P_0 - AC) \times Q_{PM0}$.

LR: Market price, P_1 is tangent to the AC curve at its minimum point. Profit-maximizing quantity is now Q_{PM1}, which is lower than Q_{PM0}. In the long run, firms only make normal profits, or zero economic profits.

Remember the following two conclusions from the analysis in Figure 4-6.

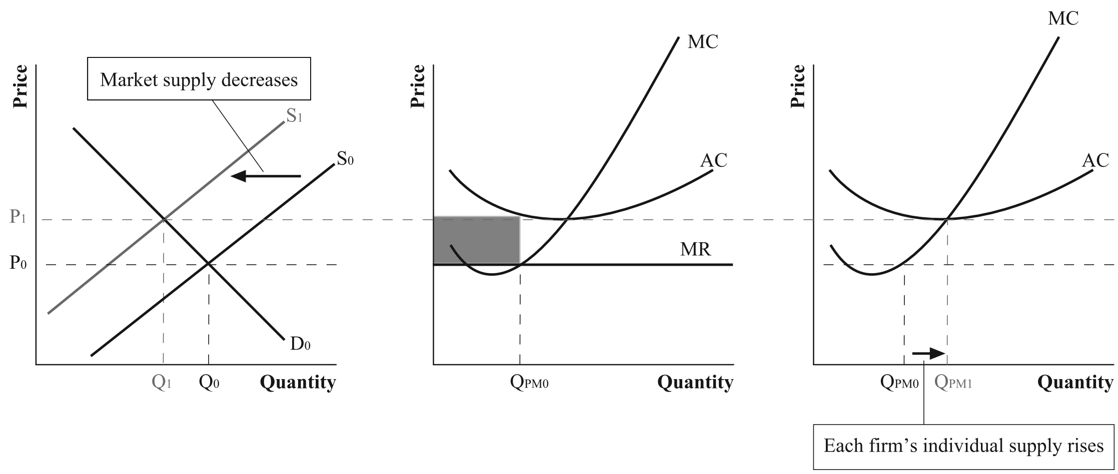
1. There are no LR economic profits in a perfectly competitive industry.
2. In the LR, price equals minimum average cost and firms make normal profit.

Economic Losses

If an industry is making economic losses, participating entrepreneurs will exit in order to make at least normal profits elsewhere. There are no barriers to exit, and every firm has an identical cost structure. When firms leave the industry, supply falls, prices rise, and eventually, economic losses are eliminated as illustrated in Figure 4-7. Remaining firms earn normal profits in the long run.

Figure 4-7: SR Losses, LR Normal Profits

4-7a. Short Run and Long Run: Industry	**4-7b. Short Run: Individual Firm**	**4-7c. Long Run: Individual Firm**



SR: S_0 and D_0 interact to set prices at P_0. **LR:** Some firms exit the industry so the industry supply gradually moves to S_1 to set LR market prices at P_1.	**SR:** Each firm produces Q_{PM0} where MR = MC. AC exceeds price at Q_{PM0} resulting in economic losses (region shaded in grey) equal to $(AC - P_0) \times Q_{PM0}$.	**LR:** Market price, P_1 is tangent to the AC curve at its minimum point. Profit-maximizing quantity is now Q_{PM1}, which is greater than Q_{PM0}. In the long run, firms only make normal profits, or zero economic profits.

Perfect competition is discussed further in the next reading.

Important takeaways from the analysis in Figure 4-7:

1. There are no long-run economic losses in a perfectly competitive industry.
2. Price equals minimum average cost in the long run and firms make normal profit.

LESSON 5: LONG RUN SUPPLY

LOS 15i: Distinguish among decreasing-cost, constant-cost, and increasing-cost industries and describe the long-run supply of each.

The Long Run Industry Supply Curve

External economies are factors outside the control of the firm that *decrease* average costs for individual firms as *industry output increases*. Examples of external economies are specialists who develop consulting practices when the number of firms in the industry (potential clients) increases. They use their experience and knowledge to help firms become more cost-effective and efficient, bringing average costs lower.

External diseconomies are factors outside the control of the firm that *increase* average costs for individual firms as *industry output increases*. An example of external diseconomies is how an increase in flights and number of airlines causes congestion at airports and results in longer waiting times and airport charges for all airlines, which increase average costs.

The LR industry supply curve shows how quantity supplied changes as market prices vary after all possible adjustments have occurred (including changes in plant size and changes in the number of firms in the industry). For example, an increase in demand in a perfectly competitive industry results in higher prices, which attracts new entrants. Over the long run, the entry of these new firms increases industry supply, which brings prices down to a level where economic profits are eliminated.

In constant-cost industries, supply increases by as much as the initial increase in demand such that prices return to their original levels in the long run. As a result, the long run supply curve is *perfectly elastic* (see Figure 5-1a).

> Examples of increasing cost industries include the petroleum, coal, and natural gas industries where long-run demand growth results in higher output prices because of the rising costs of energy production.

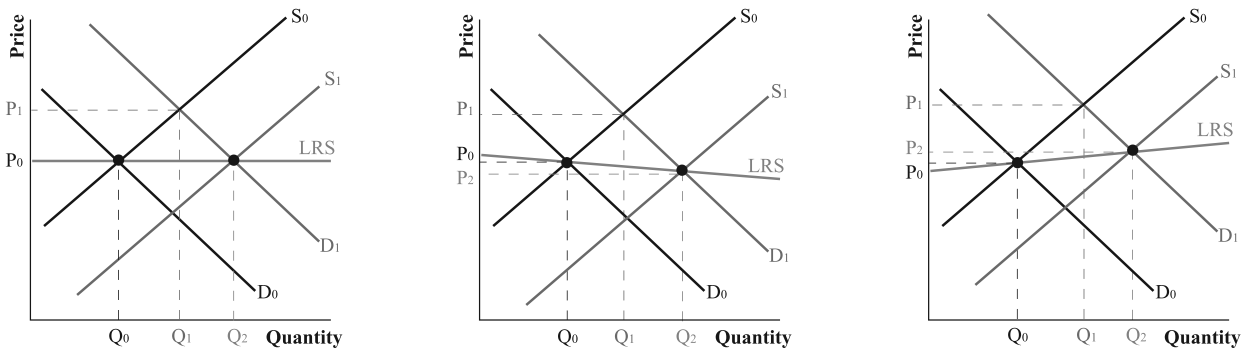
In industries with external economies (decreasing-cost industries), the presence of a larger number of firms lowers costs for all firms. Firms are able to bring down prices as they incur lower resource costs. The magnitude of shift in supply is greater than that of the initial shift in demand, and prices fall below original levels. The long-run supply curve for decreasing-cost industries is *downward sloping* (see Figure 5-1b).

In industries with external diseconomies (increasing-cost industries), an increase in demand boosts prices, but as more firms enter, average costs for all firms rise. Since the industry faces higher production costs, firms will charge a higher price for their output. Supply increases by less than the initial increase in demand. This results in prices that are higher than original levels, and a long-run supply curve that is *upward sloping* (see Figure 5-1c).

> Note that even though the industry's LR supply curve is downward sloping in a decreasing-cost industry, an individual firm's supply curve would still be upward sloping.

> Examples of decreasing-cost industries include semiconductors and personal computers, where an increase in demand has led to significantly lower prices.

Figure 5-1: Long-Run Supply



5-1a. Constant-Cost Industry

No external economies.

Short run: When demand increases to D_1, prices rise to P_1, and equilibrium quantity moves to Q_1.

Firms enter industry in search of economic profits. Therefore, supply increases to S_1, prices revert to P_0, and equilibrium quantity increases further to Q_2.

Long run: Flat LRS curve.

5-1b. Decreasing-Cost Industry

External economies.

Short run: When demand increases to D_1, prices rise to P_1, and equilibrium quantity moves to Q_1.

Firms enter industry in search of economic profits. The entry of more firms results in a decrease in costs for all firms. Therefore, the magnitude of the increase in supply is greater than that of the initial increase in demand. Prices end up at P_2, which is lower than original prices.

Long run: LRS has a negative slope.

5-1c. Increasing-Cost Industry

External diseconomies.

Short run: When demand increases to D_1, prices rise to P_1, and equilibrium quantity moves to Q_1.

Firms enter industry in search of economic profits. The entry of more firms results in an increase in costs for all firms. Therefore, the magnitude of the increase in supply is lower than that of the initial increase in demand. Prices end up at P_2, which is higher than original prices.

Long run: LRS has a positive slope.