

CHAPTER ONE: - INTRODUCTION TO ECONOMIC ANALYSIS

1.1. Scope and Methods of Economic Analysis

1.1.1. The concept of economic theory

A theory is a framework that helps us to understand the relationships between cause and effect and the implications of each type of relationships in the economic process. A theory can be valid not only by its realism but by its power to predict and explain economic phenomena.

Economic theory aims at the construction of models which describe the economic behavior of individual units such as consumers, producers, business firms and governments, and their interaction which create the economic system of a region, country or the world as a whole.

A model is a simplified representation of a real situation. It includes the main features of the real situation which it represents.

There are two main purposes for which a model is built:

- I. Analysis:- it implies the explanation of the behavior of economic units (consumers and producers)
- II. Prediction:- it implies the possibility of forecasting the effects of changes in some magnitudes in the economy under study.

1.1.2. Definition and the concept of economics

Economics is a social science which studies about efficient allocation of scarce resources so as to attain unlimited human needs. Economics is a science of choice. It studies how people choose to use scarce or limited productive resources (land, labor, equipment, technical knowledge and the like) to produce various commodities.

- *Scarcity* means that society has limited resources and therefore cannot produce all the goods and services people wish to have.
- *Efficiency* means society gets the most that it can from its scarce resources.

In sum, economics is concerned with activities that involve:

- Production
- Distribution, and
- Exchange of goods and services in the economy.

1.1.3. Branch of economics

Economics can be analyzed based on two main branches – microeconomics and macroeconomics.

- **Microeconomics** is a branch of economics that focuses on the individual parts of the economy. That is, how households and firms make decisions and how they interact in specific markets.
- **Macroeconomics** is a branch of economics that deals with the effects and consequences of the aggregate behavior of all decision making units in a certain economy. It looks at the economy as a whole. It discusses about the economy-wide phenomena, including inflation, unemployment, and economic growth.

Note: Both microeconomics and macroeconomics are complementary to each other. That is, macroeconomics cannot be studied in isolation from microeconomics.

Economics can also be analyzed from two perspectives: positive economics and normative economics.

- **Positive economics:** it is concerned with analysis of facts and attempt to describe the world as it is. It tries to answer the questions what was; what is; or what will be? It does not judge a system as good or bad, better or worse.

- **Normative economics:** It deals with the questions like, what ought to be? Or what the economy should be? It evaluates the desirability of alternative outcomes based on one's value judgments about what is good or what is bad. In this situation since normative economics is loaded with judgments, what is good for one may not be the case for the other.

1.1.4. Methods of economic analysis

There are two methods of economic analysis known as inductive method and deductive method.

7. **Inductive method** is a logical method of reaching at a correct general statement or theory based on several independent and specific correct statements. In short, it is the process of deriving a principle or theory by moving from facts to theories and from particular to general economic analysis.

2. **Deductive method** is a logical way of arriving at a particular or specific correct statement starting from a correct general statement. In short, it deals with conclusions about economic phenomena from certain fundamental assumptions or truths or axioms through a process of logical arguments. The theory may agree or disagree with the real world and we should check the validity of the theory to facts by moving from general to particular.

1.2. The Fundamental Economic Problems

1.2.1. Scarcity and the problem of choice

Road space throughout the world is becoming increasingly scarce as the demand for motor transport increases each year – what do you think are some of the best solutions to reducing the problem of congestion on our roads?

It is often said that the central purpose of economic activity is the production of goods and services to satisfy consumer's needs and wants i.e. to meet people's need for consumption both as a means of survival but also to meet their ever-growing demand for an improved lifestyle or standard of living.

The **basic economic problem** is about **scarcity** and **choice** since there are only a limited amount of resources available to produce the unlimited amount of goods and services we desire. Thus, economics is the study of how human beings make choices to use scarce resources as they seek to satisfy their unlimited wants.

Making a choice normally involves a **trade-off** - in simple terms, choosing more of one thing means giving up something else in exchange. Because wants are unlimited but resources are finite, choice is an unavoidable issue in economics. This means that scarcity implies economic choice and choice in turn implies sacrifice of the next best alternative i.e. each choice has an **opportunity cost**.

Opportunity costs: - The best alternative that we give up, or forgo, when we make a choice or decision.

Example: Education takes time. Time spent in the classroom has an opportunity cost.

As described above, the central problem that all societies face is the **scarcity of resources as compared to the unlimited wants of their members**. Thus, economic systems try to organize and allocate these scarce resources among competing ends and withstand the imbalance between limited capacity of productive resources and unlimited human wants or desires.

Therefore, all societies face the problem of having to decide:

- What goods and services to produce: It refers to the identification of what mix of goods and services to produce and what quantity of those goods and services to produce over a period of time. Does the economy uses its resources to operate more hospitals or hotels? Do we make iPod or produce more coffee?

- How best to produce goods and services: This refers to a choice between alternative technologies or techniques of production. Once the possible and required combinations of outputs are determined, the system has to determine the most suitable way of producing them. On the other way, we must decide on the best use of our scarce resources of land, labour and capital? Should school playing fields be sold off to provide more land for affordable housing? Or should it be sold off for the construction of public hospital?

- Who is to receive goods and services: it refers to a choice between alternative ways of distributing the final products. That is, who gets how much of what is produced in the economy. What is the best method of distributing products to ensure the highest level of wants and needs are met? Who will get expensive hospital treatment - and who not?

1.2.2. The production possibility frontier and opportunity cost

The production possibility frontier (PPF) which sometimes called production possibility curve (PPC) or product transformation curve (PTC) is an important tool to understand the problem of scarcity and choice. The PPF depicts the curve showing different combinations of various goods that can be produced from a given amount of resources. In other word, it represents the point at which an economy is most efficiently producing its goods and services and, therefore, allocating its resources in the best way possible. If the economy is not producing the quantities indicated by the PPF, resources are being managed inefficiently and the production of society will decline. The production possibility frontier shows there are limits to production, so an economy, to achieve efficiency, must decide what combination of goods and services can be produced.

Illustration

The PPF we use for illustration is based on the following assumptions:

XII. The economy can produce only two goods: example wine and cotton

XIII. The quantity of resources (land, labour, capital, entrepreneur) the economy is endowed is assumed fixed for a period of time.

XIV. Technology is assumed given and constant in the short run.

XV. Resources are assumed used efficiently. That is, no resources are remaining idle or no resources are misallocated when production takes place.

According to the PPF, points A, B and C - all appearing on the curve - represent the most efficient use of resources by the economy. Point X represents an inefficient use of resources, while point Y represents the goals that the economy cannot attain with its present levels of resources.

As we can see, in order for this economy to produce more wine, it must give up some of the resources it uses to produce cotton (point A). If the economy starts producing more cotton (represented by points B and C), it would have to divert resources from making wine and, consequently, it will produce less wine than it is producing at point A. As the chart shows, by moving production from point A to B, the economy must decrease wine production by a small amount in comparison to the increase in cotton output. However, if the economy moves from point B to C, wine output will be significantly reduced while the increase in cotton will be quite small. Keep in mind that A, B, and C all represents the most efficient allocation of resources for the economy; the nation must decide how to achieve the PPF and which combination to use. If more wine is in demand, the cost of increasing its output is proportional to the cost of decreasing cotton production.

Point X means that the country's resources are not being used efficiently or, more specifically, that the country is not producing enough cotton or wine given the potential of its resources. Point Y, as we mentioned above, represents an output level that is currently unreachable by this economy. However, if there was a change in **technology** whilst the level of land, labor and capital remained the same, the time required to pick cotton and grapes would be reduced. Output would increase, and the PPF would be pushed outwards. A new curve, on which Y would appear, would represent the new efficient allocation of resources.

When the PPF shifts outwards, we know there is growth in an economy. Alternatively, when the PPF shifts inwards it indicates that the economy is shrinking as a result of a decline in its most efficient allocation of resources and optimal production capability. A shrinking economy could be a result of a decrease in supplies or a deficiency in technology.

Opportunity Cost

Opportunity cost is the value of what is foregone in order to have something else. This is important to the PPF because a country will decide how to best allocate its resources according to its opportunity cost. Therefore, the previous wine/cotton example shows that if the country chooses to produce more wine than cotton, the opportunity cost is equivalent to the cost of giving up the required cotton production.

Opportunity cost of shifting production from A to B =

Let's look at another example to demonstrate how opportunity cost ensures that an individual will buy the least expensive of two similar goods when given the choice. For example, assume that an individual has a choice between two telephone services. If he or she were to buy the most expensive service, that individual may have to reduce the number of times he or she goes to the movies each month. Giving up these opportunities to go to the movies may be a cost that is too high for this person, leading him or her to choose the less expensive service.

Remember that opportunity cost is different for each individual and nation. Thus, what is valued more than something else will vary among people and countries when decisions are made about how to allocate resources.

1.2.3. Decision making units and circular flow of economic activities

The three most important decision making units are:

1. Households: - Households are economic units that provide an economy with the resources they sell in the factor market and use the money paid for them to buy goods and services from the product market.

- They are owners of resources (land, labour, capital and entrepreneurship)
- They make decisions on how to sell their resources to firms and governments.
- They make decisions on what and how much of the commodities they can buy from business firms.

2. Business firms: - A firm is a production unit that uses resources to produce goods and services and sell them to other firms, households and governments.

- They make decision on buying resources from households.
- They make decision on selling their products to households and governments.

3. Government: - Refers to an organization that has a legal and political power to exert control over individuals, business firms and markets.

- Provides social services such as defense, education, public health and infrastructural facilities.
- Tax from households and business firms is the main source of government revenue.

These three decision making units interact in product and factor markets to transact goods and services, and resources (factors of production), respectively.

• **Product market** is a market where goods and services are transacted/ exchanged. That is, a market where households and governments buy goods and services from business firms.

• **Factor market (input market)** is a market where economic units transact/exchange factor of production (inputs). In this market, owners of resources (households) sell their resources to business firms and governments.

The **circular-flow diagram** is a visual model of the economy that shows how money (Birr) flows through markets among the decision making units.

For simplicity, let's first see a two sector model where we have only households and business firms i.e. no government. In this case, therefore, we see the flow of goods and services from producers to households and a flow of resources from households to business firms.

In the following diagram the clock – wise direction shows the flow of economic resources and final goods and services (flow of inputs and outputs). Business firms sell goods and services to households in product markets (upper part of the diagram). The lower part shows, on the other hand, where households sell factors of production to business firms through factor market. However, the anti – clock wise direction indicates the flow of birr (in the form of revenue, income and spending on consumption). Firms by selling goods and services to household receive money in the form of revenue which is consumption expenditure for households in the product market. On the other hand, households by supplying their resources to firms receive income. This represents expenditure by firms to purchase factors of production which is used as an input to produce goods and services.

We have also a three sector model in which the government is involved in the economic activities. As shown in diagram below, the only difference of the three sector model from the two sector model is in that it involves government participation in the market. The government to provide public services purchase goods and service from business firms through the product market with a given amount of expenditure. On the other hand, the government also needs resources required for the provision of the services. This resource is purchased from the factor market by making payments to the resource owners (households).

The service provided by the government goes to the households and business firms. The government might also support the economy by providing income support to the households and subsidies to the business firms. At this point you might ask the source of government finance to make the expenditures, payments and additional supports to the firms and households. The main source of revenue to the government is the taxes collected from households and firms.

1.2.4. Alternative economic systems

An economic system may be defined as a network of decisions and actions of economic and political players in production, distribution and consumption process in a country. Different societies are organized through alternative economic systems. We generally distinguish two fundamentally different ways of organizing an economy. At one extreme, government makes economic decision, with those on top of the hierarchy giving economic commands to those further down the ladder. At the other extreme, decisions are made in markets, where individuals or enterprises voluntarily agree to exchange goods and services, usually through payments of money.

21. Free -market economy (Capitalism): - A free market economy is one in which individuals and private firms make the major decisions about production and consumption. A system of prices, of markets, of profits and losses, of incentives and rewards determines what, how and for whom basic economic questions. That is, in a free market economy the basic economic problem of what, how, and for whom to produce are resolved by the market mechanism. Firms resolve the economic problem of what to produce by producing goods and services that yield high profit. The choice of technology for the how to produce problem is resolved based on the least cost combination of inputs and the for whom to produce is resolved based on the owners of the means of production.

2. Command economy (Socialism): - By contrast, a command economy is one in which the government makes all important decisions about production and distributions. In short, in a command economy, the government answers the major economic questions through its ownership of resources and its power to enforce decisions. Production of goods and services are undertaken for the uses keeping in mind the needs of consumption rather than profit.

3. Mixed economy: - No contemporary society falls completely into either of command or market polar categories. Rather, all societies are mixed economies, with elements of market and command. There has never been a 100 percent market economy. The basic economic problems are resolved by a mixture of government decisions and market forces of demand and supply. However, the role of the government is intermediate in a mixed economy. Currently most of the world economies are mixed with different degrees of mixture.

4. Traditional economy: - In this type of economic system, the basic economic problems are resolved by traditional or long standing rules or customs of social behavior. Production, exchange and distribution are sanctioned by custom. Technological changes and innovations are usually constrained by tradition. In general, economic activities are treated as secondary and societies mainly focus on religious and cultural values in a traditional economy.

Review Questions

True or False Questions

25. Economic models and theories are accurate statements of reality.
2. Graphs provide a visual representation of the relationship between two variables.
3. A production-possibility frontier depicts the unlimited wants of a society.
4. When there is full employment, the decision to produce more of one good necessitates decreased production of another good.
5. There are increasing costs of production because economic resources are not equally efficient in the production of all goods and services.

Multiple Choice Items

30. In the factor market,
 - a. Households are sellers of factors of production.
 - b. Firms are sellers of goods and services.
 - c. Households are buyers of goods and services.
 - d. Firms are buyers of factors of production and sellers of goods and services.
2. Among the types of economic systems, one provides the fairest distribution of income.

- a. Mixed
- b. Command
- c. Market
- d. A and C

4. In the circular flow diagram of an economy, firms represent the ____ side of the input market while households represent the ____ side of the product market.

- a. supply/supply
- b. supply/demand
- c. demand/supply
- d. demand/demand

6. Identify the statement that is normative in nature.
- The Millennium Abay Dam enables the country to export electric power.
 - The Millennium Abay Dam is important to reduce poverty in Ethiopia.
 - The Millennium Abay Dam needs to consider environmental protection issues.
 - All are positive statements.
7. Choose the correct combination of an economic problem and its focus.
- What to produce – technique of production
 - For whom to produce – distribution of products
 - How to produce – allocation of resources
 - A and C

Work Out Items

35. Assume that a certain simplified economy produces only two goods, X and Y, with given resources and technology. The following table gives the various possible combinations of the production of the two goods (all units are measured in millions of tons).

Production Possibility	Good X	Good Y	Opportunity Cost of Good X
A	0	100	
B	2	90	
C	4	60	
D	6	20	

- Calculate the opportunity cost of the production of good X. What law does the trend in those values exhibit?
- What things are necessary for that economy to register economic growth?

CHAPTER 2: DEMAND, SUPPLY AND MARKET EQUILIBRIUM

Having covered the nature of economics in the first chapter, we now resort to a very important issue in economics. This is the issue of how free markets operate. A fundamental task of economics is to analyze the factors that determine the prices and quantities of commodities sold. For this purpose we will first discuss important concepts that help us to understand how the market system functions. The important concepts are the theory of demand and supply. The supply and demand model is designed to explain how prices are determined in a market

The purpose of this chapter is to explain what demand and supply are and show how they determine price and the quantity sold in markets. We will also show how the concepts of demand and supply reveal consumers' and producers' sensitivity to price change.

This chapter will also explain the concept of elasticity which is a measure of the sensitivity of quantity demanded and quantity supplied.

2.1. Theory of demand

In this discussion, we will explain what demand is and what factors influence it.

2.1.1. Definition and Basic Concepts of Demand

In our day-to-day life we use the word 'demand' in a loose sense to mean a desire of a person to purchase a commodity or service. But in economics it has a specific meaning. Demand implies more than a mere desire to purchase a commodity. It states that the consumer must be willing and able to purchase the commodity, which he desires. His desire should be backed by his purchasing power. A poor person is willing to buy a car; it has no significance, since he has no ability to pay for it. On the other hand, if his desire to buy the car is backed by the purchasing power then this constitutes demand. Demand, thus, means the desire of the consumer for a commodity backed by purchasing power. These two factors are essential. If a consumer is willing to buy but is not able to pay, his desire will not become demand. Similarly, if the consumer has the ability to pay but is not willing to pay, his desire will not be called demand.

More specifically, demand refers to various quantities of a commodity or service that a consumer would purchase at a given time in a market at various prices. In deriving the demand of a commodity, it is assumed that other things such as consumer's income, tastes, prices of interrelated goods, etc, remain unchanged (*ceteris paribus*). The quantity demand of a particular commodity depends on the price of that commodity.

The relationship that exists between price and the amount of a commodity purchased can be represented either by an equation or a table (schedule) or a curve.

2.1.2. The Law of Demand

Law of demand tells us the functional relationship between the price of a commodity and its quantity demanded in the market. The law of demand states that all other factors remaining the same, the price of a good and its quantity demanded have inverse relation, i.e., as price rises quantity demanded falls and vice versa.

Why do you think there is an inverse relation between price and quantity demand for most goods?

The following are the most common explanations to this law.

i. The law of diminishing marginal utility of goods

This explanation was given by classical economists. As the additional utility derived from consuming successive units of a good decrease, the consumer will be willing to take more at a falling price.

ii. Substitution effect

This is because consumers are rational and they always tend to substitute cheap product for an expensive product.

iii. Income Effect

This is because the price change has an effect on the purchasing power of the consumer and therefore he can buy less when price increases and more when price decreases.

Exceptions to the law of demand

The law of demand may not apply to some cases. Some of these are discussed as follows.

Giffen's Paradox

As the law of demand states if the price increases then the demand should decrease. But, sometimes it may rise. In other words, sometimes people may buy more when the price of a good increases. This was first observed by Sir Robert Giffen.

Prestigious goods

Sometimes people purchase certain goods as their possession confers a higher social status on them. For example diamonds and precious stones are purchased by rich people to maintain high prestige in the society without caring for the high price of goods.

Speculation

Sometimes, the price of a commodity might be increasing and it is expected to increase still further. The consumers will buy more of the commodity at the higher prices than they did at lower price. The increase in price may not be accompanied by a decrease in its demand, which is contrary to the law of demand.

2.1.3. Demand Equation, Schedule and Curve

The relationship between the price of a good and its quantity demanded, all other things remaining the same can be represented using an equation where quantity demanded will be the dependent variable and price will be the exogenous variable as given below.

$D_x = f(P_x)$ where D_x = quantity demanded

P_x = price of commodity

The specific function can take any form from linear to non-linear functions of various forms. Demand schedule is a tabular way of representing the relation between price and quantity demanded.

A demand schedule can be constructed for any commodity if the list of prices and quantities purchased at those prices are known. An individual demand schedule is a list of the various quantities of a commodity, which an individual consumer purchases at various levels of prices in the market. A demand schedule states the relationship between two variables of price and quantity demanded in a table form.

Demand schedule for oranges

Price per dozen (Birr)	Quantity demanded in Kgs
1	15
2	12
3	10
4	7

Demand curve: graphical representation of demand

In the above diagram prices of oranges are given on 'OY' axis and demand on 'OX' axis. When the price per Kilogram is birr 1 only, 15 kilograms are demanded. If we plot the data as above, you may notice that if the price falls down demand increases and vice-versa.

2.1.4. Individual and Market Demand

An individual's quantity demanded of any good is the total amount that individual would choose to buy at a particular price. While, the market quantity demanded of any good is the total amount that all buyers in the market would decide to buy at a particular price.

For example, the demand for 'teff' by a single household living in Addis Ababa is an individual demand while the total demand for 'teff' in Addis Ababa is the market demand.

Market demand is simply the horizontal sum of individual demand.

2.1.5. Determinant of Demand

The demand for a product is influenced by many factors. Some of these factors are:

- XXXVI. Price of the product
- XXXVII. Taste or preference of consumers
- XXXVIII. Income of the consumers
- XXXIX. Price of related goods
- XL. Consumers expectation of income and price
- XLI. Number of buyers in the market

When we state the law of demand, we kept all the factors to remain constant except the price of the good.

A change in any of the above listed factor except the price of the good will change the demand. While a change in the price, other factors remain constant will bring change in quantity demanded. A change in demand will shift the demand curve from its original location for these reason those factors listed from II to VI are called demand shifters. While a change in own price is only a movement along the same demand curve.

Changes in demand: a change in any determinant of demand—except for the good's price—causes the demand curve to shift. We call this a change in demand. If buyers choose to purchase more at any price, the demand curve shifts rightward—an increase in demand. If buyers

choose to purchase less at any price, the demand curve shifts leftward—a decrease in demand.

Now let us examine how each factor affect demand

I) Taste or preference

When the taste of a consumer changes in favor of a good, her demand will increase the opposite change will decrease demand.

II) Income of the consumer

Goods are classified into two depending on how a change in income affects their demand.

Normal Goods: are goods whose demand increases, as income increase. While inferior goods are those whose demand is inversely related with income, i.e, as income increases, the demand for these goods decrease.

In general, inferior goods are poor quality goods with relatively lower price and buyers of such goods are expected to shift to better quality goods as their income increases. However, the classification of goods into normal and inferior is subjective and it is usually dependent of the socio-economic development of the nation.

III) Price of related goods

Related goods are those goods which have some form of relation. Two goods are said to be related if a change in the price of one good, affect the demand of another good.

There are two forms of relation between goods. These are substitute and complimentary goods.

Substitute goods are goods which satisfy the same desire of the consumer. For example, tea and coffee or Pepsi and Coca-Cola. If two goods are substitutes, then price of one and the demand of the other are directly related.

Complimentary goods, on the other hand are those which are jointly consumed. For example, car and fuel or tea and sugar. If two goods are complements, then price of one and the demand of the other are inversely related.

IV) Consumer expectation of income and price

Higher price expectation will increase demand while a lower future price expectation will decrease the demand for the good.

Higher income expectation will increase demand while a lower future income expectation will decrease the demand for the good.

V) Number of buyer in the market

Since market demand is the horizontal sum of individual demand, an increase in the number of buyers will increase demand while a decrease in the number of buyers will decrease demand.

2.2. Theory of Supply

In this discussion, we will explain what supply is and what factors influence it.

2.2.1. Definition and Basic Concepts of Supply

Supply means the quantity of a commodity that a seller is willing and able to offer for sale. If the price goes up, the producer will offer more for sale. But, if the price goes down, he will be unwilling to sell and will offer less to sell. Hence, supply varies with a change in price and supply is always made at a price. Just as demand implies willingness and ability to pay, "Supply implies willingness and ability to deliver the goods". There is a difference between the terms 'stock' and supply. Stock constitutes potential supply. But supply means the quantity actually offered for sale at a certain price. Stock is that which can be offered for sale if the conditions are favorable.

More specifically, supply refers to various quantities of a commodity or service that a producer would supply at a given time in a market at various prices. In deriving the supply of a commodity, it is assumed that other things such as resource price, technology, prices of related goods and all other factors remain unchanged (*ceteris paribus*). The quantity supplied of a particular commodity depends on the price of that commodity.

The relationship that exists between price and the amount of a commodity supplied can be represented either by an equation or a table (schedule) or a curve.

2.2.2. The Law of Supply

Law of supply tells us the functional relationship between the price of a commodity and its quantity supplied in the market. The law of supply states that all other factors remaining the same, the price of a good and its quantity supplied have direct relation, i.e., as price rises quantity supplied also rises and vice versa.

Why do you think there is a direct relation between price and quantity supplied of goods?

This is because price is an incentive for the producer and higher price represents higher profit.

Exceptions to the law of supply

42. When prices are expected to fall greater, sellers sell more in order to clear their stocks. This will happen in short-run.
2. In the long run the supply is influenced by changes in costs, which are in turn affected, by changes in technology.
3. Changes in habits, tastes, fashions, weather etc. also affect the supplies of commodities.

2.2.3. Supply Equation, Schedule and Curve

The relationship between the price of a good and its quantity supplied, all other things remaining the same can be represented using an equation where quantity supplied will be the dependent variable and price will be the exogenous variable as given below.

$$S_x = f(P_x) \dots \text{where } S_x = \text{quantity demanded}$$

P_x = price of the commodity

The specific function can take any form from linear to non-linear functions of various forms.

Supply schedule is a tabular way of representing the relation between price and quantity supplied.

A supply schedule can be constructed for any commodity if the list of prices and quantities supplied at those prices are known. An individual supply schedule is a list of the various quantities of a commodity which an individual producer supplies at various levels of prices in the market. A supply schedule states the relationship between two variables of price and quantity supplied in a table form.

The following table shows a hypothetical supply schedule for mangoes.

Table: Supply schedule

Price per kgs (In Birr).	Quantity (in Kgs)
--------------------------	-------------------

- a. 400
- b. 300

3	200
---	-----

- c. 100

1	50
---	----

In the above schedule you may notice that as price falls, fewer mangos are offered for sales and as price rises, the seller is prepared to sell more of them. The simple explanation here is that the higher the price of commodity, the greater are the profits that can be earned and thus, the greater the incentive to produce and offer the commodity for sale.

This means that as the prices rises, more is offered for sale and vice-versa. This supply curve is related to the cost structure of the firm.

2.2.4. Individual and Market Supply

An individual's quantity supplied of any good is the total amount that individual would choose to sale at a particular price. While, the market quantity supplied of any good is the total amount that all producers in the market would decide to sale at a particular price.

For example, the supply of mangoes by a supper in a market is an individual supply while the total mango offered by all suppliers in the market is the market supply of mangos.

Market supply is simply the horizontal sum of individual supply.

2.2.5 Determinant of Supply

The supply of a product is influenced by many factors. Some of these factors are:

- I) Price of the product
- II) Resource price
- III) Technology
- IV) Price of related goods
- V) Price expectation of the supplier
- VI) Tax and Subsidy
- VII) Number of sellers in the market

When we state the law of supply, we kept all the factors to remain constant except the price of the good.

A change in any of the above listed factor except the price of the good will change the supply. While a change in the price, if the other factors remain constant will bring change in quantity supplied. A change in supply will shift the supply curve from its original location for these reason those factors listed from II to VII are called supply shifters. While a change in own price is only a movement along the same supply curve.

Changes in supply is the result of a change in any determinant of supply —except for the good's price and it will bring a shift in the supply curve. We call this a change in supply. If suppliers choose to supply more at any price, the supply curve shifts down ward—an increase in supply. If suppliers choose to supply less at any price, the supply curve shifts up ward—a decrease in supply.

Now let us examine how each factor affect supply.

I) Resource price

Resource price is one of the most important factors that influence supply. Since resource price directly affect the cost of the firm and hence its profit. Therefore, an increase in resource price is likely to decrease supply since the profit of the firm declines with the higher resource prices and vice versa.

II) Technology

Technology is another factor which influences the supply of a good. Technological advancement will increase the supply of the good.

III) Price of related goods

Goods may have some relations in production. This relation, like in the demand side, can be either substitutability or complementarity.

Substitute goods in production involve two goods competing for the given resource of the supplier. For example a farmer may have the option of producing either maize or Chat on her plot of land. A rise in the price of Chat will encourage the farmer to use her limited resource for producing more chat and hence the supply of maize (food) will be decreased.

Complimentary goods, on the other hand are those which are produced together i.e., as we go for the production of a good another product is also produced in the same line of production. For example if the price of meat increases, the supply of animal skins will also increase. This is because as the suppliers of meat try to increase their supply because of the higher price for meat, they slaughter more animals making more skins available in the market.

IV) Suppliers expectation of price

In general, when suppliers expect higher price in the future they tend to decrease the supply of their product now in order to sale it at higher expected price in the future.

V) Tax and subsidy

Tax and subsidies also affect the supply of a good.

An increase in tax rate is likely to decrease supply since the supplier faces higher cost of production. On the other hand, provision of subsidy encourages the suppliers to increase their supply.

VI) Number of sellers in the market

Since market supply is the horizontal sum of individual supply, an increase in the number of sellers will increase supply while a decrease in the number of sellers will decrease supply.

2.3. Market Equilibrium

In this section we will bring the discussion of demand and supply to see how they interact to set the market equilibrium and market variables. But first let us examine briefly the various types of equilibrium analysis.

2.3.1. Types of Equilibrium Analysis

The term equilibrium has often to be used in economic analysis. In fact, Modern Economics is sometimes called equilibrium analysis. Equilibrium means a state of balance. When forces acting in opposite directions are exactly equal, the object on which they are acting is said to be in a state of equilibrium.

Equilibrium analysis in economics may take the following different forms

A) Partial Equilibrium: Partial equilibrium analysis is the analysis of an equilibrium position for a sector of the economy or for one or several partial groups of the economic unit corresponding to a particular set of data. This analysis excludes certain variables and relationship from the totality and studies only a few selected variables at a time. In other words, this method considers the changes in one or two variables keeping all others constant, i.e., *ceteris paribus* (others remaining the same). The *ceteris paribus* is the always a very important part of partial equilibrium analysis.

The equilibrium of a single consumer, a single producer, a single firm and a single industry are examples of partial equilibrium analysis.

B) General Equilibrium Analysis: The partial equilibrium analysis studies the relationship between only selected few variables, keeping others unchanged. Whereas the general equilibrium analysis enables us to study the behaviour of economic variables taking full account of the interaction between those variables and the rest of the economy. In partial equilibrium analysis, the determination of the price of a good is simplified by just looking at the price of one good and assuming that the prices of all other goods remain constant.

Partial equilibrium analysis is what we are going to discuss in this chapter.

2.3.2. Market Equilibrium and its Determination

Market equilibrium is a state of the market in which market forces of demand and supply are in balance and as a result the market price and the amount of the good exchanged tend to be stable.

The buyers and sellers in the market bargain about goods and services. They agree to purchase and sell goods and services at a certain price. In this manner the price is determined by the interaction of buyers and sellers. In other words, demand and supply determine prices.

Buyers in the market follow the law of demand. The Law of demand reveals that when the price falls the demand increases and when the price raises the demand decreases. The sellers follow the law of supply. According to this law, when the price rises the supply is increased and when the price falls the supply is decreased. Thus, demand and supply move in opposite directions. Price is determined by the interaction of demand and supply and the price at which demand and supply are equal is known as the equilibrium price. Equilibrium quantity is the quantity supplied and the quantity demanded at equilibrium price. If the price is more or less than the equilibrium price, the equilibrium output is disturbed. But ultimately the quantity demanded and the quantity supplied will be balanced at some equilibrium price.

We can explain this process of equilibrium adjustment with the help of schedule and figure. In the following table the quantity demanded and supplied is shown.

Table:- Quantity demanded and Supplied (in kgs)

Price (Birr)	Quantity	Supplied	Shortage (-) or Surplus (+)
2	200	50	-150 (\downarrow)
3	150	75	-75 (\downarrow)
4	130	100	-30 (\downarrow)
5	110	110	0
6	90	120	+30 (\uparrow)
7	80	130	+50 (\uparrow)

When the price of mango is Birr 2 per Kg., the demand for them will be 200 kgs but the quantity supplied will be 50 Kg. At this price of Birr two the market will have a shortage of 150 kg and this will not make the market stable. The effect of a shortage in a free market is always to increase the price. If the price rises to Birr 3 the supply increases to 75 Kgs, whereas the demand falls to 150 Kg. This higher price has reduced but has not eliminated the shortage so price has to decrease further. If the price is again raised, the supply will increase to 100Kg, and the demand will be further reduced to 130Kg. Again, when the price goes up to Birr 5 per Kg., the quantity demanded or supplied will be 110 Kg. At a price of Birr 5 per Kg, the buyers are ready to purchase 110 Kgs and the sellers are ready to offer 110Kgs. This is the equilibrium price and the quantity supplied at this price is called the equilibrium quantity, i.e., 110 Kgs. Once the equilibrium price is determined, there is no change from this price as this satisfies both the consumers and the producers. If, at any time, the price is more or less than Birr 5 the forces of demand and supply will adjust it back to Birr 5. For instance, if the price increases to Birr 6 the quantity demanded will decrease, but the quantity supplied will increase. At Birr 6, the demand will be 90 Kgs and the supply will be 120Kgs. When the supply is greater than the demand, the seller has to reduce the price to Birr 5. As a result, the supply will fall to 110 Kg. and demand will also increase to 110 Kgs. Thus, the equilibrium price is re-established. On the contrary, if the price decreases to Birr 4 the demand will increase to 130 Kgs and the supply will fall to 100Kgs. Here supply

will be less than the demand. Decreased supply and an increased demand will lead to a rise in the price to Birr 5- where the demand and supply will be in equilibrium. In this manner the equilibrium will be maintained again.

In the Figure below DD is the demand curve and SS is the supply curve. These two curves intersect each other at point E. The point, at which demand and supply curve intersect is called the equilibrium point. OP is the equilibrium price at which OM, the equilibrium quantity is demanded and supplied. If demand is less than supply, every seller will try to sell his quantity of the product first by reducing the price a little. Sellers compete among themselves to bring down the price to the equilibrium level. Thus demand and supply determine the equilibrium price. Once it is established, any disturbance from this equilibrium level will be restored by the forces of demand and supply.

Illustration

If the demand and supply functions are given in equation form as below

$$X_d = 100 - 2P \text{ and } X_s = -20 + 4P$$

Attempt the following questions.

- A) What is the equilibrium price and quantity?
- B) Show the equilibrium diagrammatically.
- C) If the price of the good is fixed at P=25, will there be a shortage or a surplus? How large is it?

2.3.3. Effects of Shift in Demand and Supply on Equilibrium

Given demand and supply the equilibrium price and quantity are stable. However, when these market forces change what will happen to the equilibrium price and quantity.

Changes in demand and supply bring about changes in the equilibrium price level and the equilibrium quantity. When there is a change in either the demand or the supply or both, a new equilibrium price is established. Thus, whenever change in either demand or supply takes place, a new equilibrium price is determined.

I) Effects of Changes in Demand

There are some conditions, which bring about changes in demand. A change in incomes, tastes, prices of substitute goods and preferences of consumers will lead to a change in demand. The figure below shows the effects of a change in demand and the resultant equilibrium price and quantity. DD is the demand curve and SS is the supply curve. The supply curve remains the same. Only demand changes.

DD and SS curves intersect at point E and the quantity demanded and supplied is OM at OP equilibrium price. Given the supply, if the demand increases the demand curve will shift upward to the right. Due to a change in demand, the demand curve D₁D₁ intersects SS supply curve at point E₁. The equilibrium price is increased from OP to OP₁ and the equilibrium quantity from OM to OM₁. On the other hand, if demand falls, the demand curve shifts downwards to the left. Due to a change in demand, the curve D₂D₂ intersects the supply curve SS at point E₂. The equilibrium price is OP₂ which has decreased from OP to OP₂ and the equilibrium quantity from OM to OM₂. Supply being given, a decrease in demand reduces both the price and the quantity and vice versa.

II) Effects of Changes in supply

Changes in supply are brought by changes in technical knowledge and factor prices. The following graph explains the effects of changes in supply. Demand DD being given, the supply curve is SS.

SS and DD intersect at point E, where supply and demand are equal at OM quantity at OP equilibrium price. Given the demand, if the supply increases, the supply curve shifts to the right S₁S₁. The new supply curve, which intersects DD curve at E₁, reduces the equilibrium price from OP to OP₁ and increases the equilibrium quantity from OM to OM₁. On the contrary, when the supply falls, the supply curve moves to the left (S₂S₂) and cuts the DD curve at point E₂ at which the equilibrium price is raised from OP to OP₂ reducing the equilibrium quantity from OM to OM₂.

III) Effects of Combined Changes in Demand and Supply

Now we shall study the effects of combined increases and/or decrease in demand and supply.

When both demand and supply increases, the quantity of the product will increase definitely. But it is not certain whether the price will rise or fall. If an increase in demand is more than an increase in supply, then the price goes up. On the other hand, if an increase in supply is more than an increase in demand, the price falls but the quantity increases. If the increase in demand and supply is same, then the price remains the same.

When demand and supply decline, the quantity decreases. But the change in price will depend upon the relative fall in demand and supply. When the fall in demand is more than the fall in supply, the price will decrease. On the other hand, when the fall in supply is more than the fall in demand, the price will rise. If both demand and supply decline in the same ratio, there is no change in the equilibrium price, but the quantity decreases.

The demand and supply might change in opposite directions. Given an increase in demand and decrease in supply, if the increase in demand is greater than the decrease in supply the equilibrium price and quantity will rise. However, if the decrease in supply is greater than the increase in demand while price rise, equilibrium quantity will fall. If the increase in demand is exactly equal to the decrease in supply, then the price rises, but the equilibrium quantity remains the same. The reverse will be true for a decrease in demand and an increase in supply.

Thus the price is determined at the point where both demand and supply cut each other. The only really accurate answer to the question whether it is supply or demand, which determines price, is that it is in fact both. At times it will seem that one is more important than the other, for one will be active and the other passive. For example, if demand remains constant but supply conditions vary, it is demand, which is passive and supply, active. But neither is more or less important than the other in determining price.

Exercise

Show graphically the effect a simultaneous decrease in demand and an increase in supply on equilibrium price and quantity.

2.3.4. Effects of Fixing Prices

If market forces are left to operate freely they can eliminate any surplus and shortage and the market will restore its equilibrium. However, for many reasons governments of different nations interfere with the free operation of the market by fixing the price of the good by themselves. This is what we call fixing price in a market.

Price fixing can take two forms. These are price ceiling and price floor.

A price ceiling is fixing a maximum price and sellers are not allowed to charge above the fixed price. In this case if the price is fixed below the equilibrium price, there will be excess demand (shortage) in the market as the lower price discourages the suppliers to bring the supply to the level of demand. In this situation, customers will not get all the quantity they want to buy. Therefore, a distribution mechanism other than market price is necessary.

A price floor is a minimum price below which the product is not allowed to be sold. If the price is set above the equilibrium price it will create surplus in the market unless production is limited to the level of demand using some compensation schemes. Otherwise, the government should be ready to buy the excess amount of the product due to the higher price.

2.4. Elasticity

In microeconomics elasticity is used as a method to measure the degree of dependent variables responsiveness to changes in an independent variable. Elasticity has various applications most of them are interesting and useful in the study of consumption demand.

In this section we will discuss the concept of different types of elasticities and their relevance in market.

2.4.1. Price, Income and Cross –Price Elasticity of Demand

Elasticity –means responsiveness of demand to the change in variable that affect demand. It is a measurement of the percentage of responsiveness of a quantity demanded to a percentage change in an independent variable.

There are as many elasticity of demand as its determinants. The most important of these elasticity are:

1. Price elasticity of demand
2. Income elasticity of demand
3. Cross-price elasticity of demand

1. Price Elasticity of Demand

A) Meaning of price elasticity of demand

It is a measure of the responsiveness of demand to changes in the commodity's own price. It can also be expressed as the ratio of a relative change in quantity to a relative change in price. The formula for calculating the price elasticity of demand is:

$$E_p =$$

$$E_p =$$

Where E_p stands for price elasticity of demand

If the percentages are known quantities, then the numerical size of E can be easily calculated. Let us suppose that the percentage increase in the quantity is 3 and the percentage fall in the price is 1 % then

$$E_p = -3.$$

Interpretation: For a 1% change in price, there is 3% change in quantity demanded. Price elasticity of demand is a unit less negative number.

B) Measuring price elasticity of demand

There are two types of price elasticity values. They are point and arc elasticity.

I) Point Price Elasticity of Demand

This is calculated to find elasticity at a given point. In this method, we take a straight-line demand curve joining the two axis, and measure the elasticity between two points Q and Q₁ which are assumed to be intimately close to each other.

$Ed =$

The percentage change in quantity demanded can be found out thus:

=

The percentage change in price can be found out thus:

=

$Ed =$

On a straight-line demand curve we can make use of this formula to find out the price elasticity at any particular point. We can find out numerical elasticities also on different points of the demand curve with the help of the above formula. It should be remembered that the point elasticity of demand on a straight line is different at every point.

II) Arc price elasticity of demand

The main drawback of the point method is that it is applicable only when we possess information about even the slight changes in the price and the quantity demanded of the commodity. But in practice, we do not possess such information about minute changes. We may possess demand schedules in which there are big gaps in price as well as the quantity demanded. In such cases, therefore is an alternative method known as arc method of elasticity measurement. In this method the midpoints between the old and the new data in the case of both price and quantity demanded are used. It studies a portion or a segment of the demand curve between the two points. An arc is a portion of a curve line, hence, a portion or segment of a demand curve. The formula for measuring arc elasticity is given below.

$Ed =$

Symbolically, the formula may be expressed thus:

$Ed =$

Here, Q = Original Quantity demanded

Q_1 = New quantity after change in price

P = Original price before change

P_1 = New price after change

We can take a numerical example to illustrate arc elasticity. Suppose that the price of a commodity is Br. 5 and the quantity demanded at that price is 100 units of a commodity. Now assume that the price of the commodity falls to Br. 4 and the quantity demanded rises to 110 units. In terms of the above formula, arc elasticity, then, will be

$Ed =$

=

= -

=

C) Ranges of Price elasticity of Demand

The price elasticity value can be categorized into five ranges. These are

- **Elastic demand**

This is when numerical value of price elasticity of demand is greater than one.

For example: if price elasticity is -2 then $-2 / =2 > 1$. Therefore, this is elastic demand.

For demand to be elastic a given change in price should lead to a more than proportionate change in quantity demanded.

- **Inelastic demand**

This is when numerical value of price elasticity of demand is less than one.

For example: if price elasticity is -0.7 then $-0.7 / =0.7 < 1$. Therefore, this is inelastic demand.

For demand to be elastic a given change in price should lead to a less than proportionate change in quantity demanded.

- **Unitary elastic demand**

This is when numerical value of price elasticity of demand is equal to one.

For example: if price elasticity is -1 then $-1 / =1 = 1$. Therefore, this is unitary elastic demand.

For demand to be elastic a given change in price should lead to an equal proportionate change in quantity demanded.

- **Perfectly elastic demand**

- **Perfectly inelastic demand**

D) Elasticity and total revenue

Producers of goods and services—doctors, bakers, theater owners, manufacturers, and others—can use price elasticity of demand to predict how a price change will affect their total sales revenue. And government policy makers can and do use demand elasticities to price many government services, to make tax policy, and to design programs to help the needy.

Suppose you want to use price as an instrument to increase your revenue. What would you do? Will you increase the price? This may not be the right decision, since your higher price will be accompanied by a fall in sale and therefore, you may not achieve your objective. Can you imagine a price cut to be the right decision to increase sale revenue?

To use price effectively to increase our sale revenue we must first calculate the price elasticity of our product.

Where demand is price inelastic, total revenue moves in the same direction as price. When demand is elastic, total revenue from sales moves in the opposite direction from price. Finally, where demand is unitary elastic, total revenue remains the same as price changes.

Price elasticity of xy company product is greater than 2. The marketing manager of the company should increase or decrease price for the product of the company in order to increase their sale revenue?

E) Determinants of price elasticity of demand

The basic determinants of the elasticity of demand of a commodity with respect to its own price are:

- L. The availability of substitutes: the demand for a commodity is more elastic if there are close substitutes for it.
- LI. The nature of the need that the commodity satisfies: in general, luxury goods are price elastic while necessities are price inelastic.
- LII. The time period: demand is more elastic in the long run.
- LIII. The number of uses to which commodity can be put: the more the possible uses of commodity, the lower its price elasticity will be,
- LIV. The proportion of income spent on the particular commodity
- LV. The higher the proportion of income spent on the commodity, the higher the price elasticity of demand

2. Income Elasticity of Demand

The responsiveness of demand to changes in income is termed as income elasticity of demand. It can also be expressed as the proportionate change in the quantity demanded resulting from a proportionate change in income.

$$E_y =$$

Symbolically, we may write

$$E_y =$$

For normal goods income elasticity is positive. Some of the writers have used income elasticity in order to classify goods into "luxuries" and ":" necessities". A commodity is considered to be a luxury if its income elasticity is greater than unity. A commodity is necessity is its income elasticity is small.

3. Cross Price Elasticity of Demand

It is the responsiveness of demand to change in the price of other commodities. It can also be defined as the proportionate change in the quantity demanded of X resulting from a proportionate change in the price of Y.

$$E_{xy} =$$

Cross elasticity may vary from minus infinity to plus infinity. Complementary goods will have negative cross elasticities and substitute goods will have positive cross elasticities. The main determinant of the cross elasticity is the nature of the commodities relative to their uses. If two commodities can satisfy equally well the same need, the cross elasticity is high and vice-versa.

2.4.2. Price Elasticity of Supply

It is the degree of responsiveness of the supply to change in price. It may be defined as the percentage change in quantity supplied divided by the percentage change in price. Elasticity of supply can be measured with the help of the following formula.

$$E_s = \frac{\text{Change in amount supplied}}{\text{Original amount supplied}} \quad \frac{\text{Change in price}}{\text{Original price}}$$

$$=$$

Where 'Q' refers to the quantity supplied and 'p' to the price and represents change. The supply is elastic when with a small change on price there is great change in supply. It is inelastic or less elastic when a great change in price induces only a slight change in supply. If the supply is perfectly inelastic, it will be represented by a vertical line shown as below.

Like demand price elasticity of supply can be elastic, inelastic, unitary elastic, perfectly elastic or perfectly inelastic.

The above chart represents zero elasticity in which the quantity supplied does not change as price changes. If supply is perfectly elastic it will be represented by a horizontal straight line as in second diagram. At 'OP' price the supply elasticity is infinite because nothing at all is supplied at the lower prices. Producers would supply any quantity demanded at the price.

Review Question

56. Why does the quantity of salt demanded tend to be unresponsive to changes in its price?
2. Why is the quantity of education demanded in private universities much more responsive than salt is to changes in price?
3. Summarize the relationship between price elasticity, changes in price, and changes in total expenditure.
4. To get the market demand curve for a product, why do we add individual demand curves horizontally rather than vertically?
5. The market for lemonade has 10 potential consumers, each having an individual demand curve $P = 101 - 10Qi$, where P is price in dollars per cup and Qi is the number of cups demanded per week by the i th consumer. Find the market demand curve using algebra. Draw an individual demand curve and the market demand curve. What is the quantity demanded by each consumer and in the market as a whole when lemonade is priced at $P = \$1/\text{cup}$?
6. For the demand curve $P = 60 - 0.5Q$, find the elasticity at $P=10$.
7. If the demand curve shifts parallel to the right, what happens to the elasticity at $P = 10$?
8. Consider the demand curve $Q = 100 - 50P$. Draw the demand curve and indicate which portion of the curve is elastic, which portion is inelastic, and which portion is unit elastic.
9. Suppose that at a price of \$400, 300 tickets are demanded to fly from Ithaca, New York, to Los Angeles, California. Now the price rises to \$600, and 280 tickets are demanded. Assuming the demand for tickets is linear; find the price elasticities at the quantity-price pairs (300, 400) and (280, 600).
10. The demand for schedule for beer is $Xd=25-P$, where Xd is the quantity demanded of beer in millions of barrels per year and P is price in dollars per barrel.
 - a. If the supply curve for beer is $Xs = -20 + 4P$, what is the equilibrium price of a barrel of beer?
 - b. Calculate and interpret price elasticity of demand and supply at the equilibrium point.
 - c. What would the effect on the price of a barrel of beer if a tax of \$ 4 per barrel is imposed by the government?
 - d. How much revenue does the government collect?
11. Suppose in a market there are 30,000 consumers of corn each have identical demand function given by $Qd=25-2P$ and 10,000 producers each with identical supply curve given by $Qs= -15+6P$
 - a. Find the market equilibrium price and quantity?
 - b. What would be the state of the market if market price was fixed at Birr 7 per unit?
 - c. Calculate and interpret price elasticity of demand at the equilibrium point.
 - d. Compute total revenue at market equilibrium.
12. Based on the following table which indicates expenditure of the household on a commodity, answer the questions that follow. (The price of the good is Br.10)

Income	Quantity Demanded (units / month)
(Br. / month)	

Income	Quantity Demanded (units / month)
(Br. / month)	

10,000	50
20,000	60
30,000	70
40,000	80
50,000	90

a. Calculate income elasticity of demand, if income increases from Br.10, 000 to Br. 20,000 and if income increases from Br. 40, 000 to Br. 50,000.

- b. Is this a normal or an inferior or a luxury good? Justify.
 - c. Does the proportion of household income spent on this good increase or decrease as income increases? .Why?
13. When price of tea in local café rises from Br. 10 to 15 per cup, demand for coffee rises from 3000 cups to 5000 cups a day despite no change in coffee prices.
- a. Determine cross price elasticity.
 - b. Based on the result, what kind of relation exists between the two goods?

Chapter Three: Consumer Behavior

Introduction

In chapter two, we have discussed about the demand side of the market .The theory of demand helps us to explain how changes in price, taste, income etc influence the consumption pattern of consumers. Hence, in essence, it is about the behavior of buyers. In this sense, this chapter is a continuation of the previous chapter two.

In our day –to- day life, we buy different goods and services for consumption. As consumer, we act to derive satisfaction by using goods and services .But, have ever thought of how your mother or any other person whom you know decides to buy those consumption goods and services? This is what we intend to discuss in this chapter, we want to understand their behavior in detail.

Consumer Theory

Consumer theory is to demand as producer theory is to supply. The major difference is that producer theory assumes that sellers are motivated by profit, and profit is something that one can usually directly measure. Moreover, the costs that enter into profit arise from physical properties of the production process – how many coffee cups come from the coffee cup manufacturing plant? In contrast, consumer theory is based on what people like, so it begins with something that we can't directly measure, but must infer. That is, consumer theory is based on the premise that we can infer what people like from the choices they make.

Consumer behavior can be best understood in three ways. First, by examining consumer's preference, specifically, we need a practical way to describe how people prefer one good to another thing. Second step, we must account in to the fact that consumers face budget constraints – they have limited incomes that restrict the quantities of goods they can buy. Third step, putting consumer preference and budget constraints together to determine consumer choice. In other word, given their preference and limited income, what combinations of goods will consumers buy to maximize their satisfactions?

3.1 Consumer preference

Originally, preferences were defined in terms of utility: to say a bundle (X_1, X_2) was preferred to a bundle (Y_1, Y_2) meant that the X -bundle had a higher utility than the Y -bundle. But, now we tend to think of things the other way around. The preferences of the consumer are the fundamental description that is useful for analyzing choice, and utility is simply a way of describing preferences.

Consumption bundles (consumption set): We call the objects of consumer choice consumption bundles (market basket). This is a complete list of the goods and services that are involved in the choice problem that we are investigating i.e. a collection of one or more commodities. . The word "complete" deserves emphasis: when you analyze a consumer's choice problem, make sure that you include all of the appropriate goods in the definition of the consumption bundle.

Basic assumptions of preference

Complete: We assume that any two bundles can be compared. That is, given any X -bundle and any Y -bundle, we assume that $(X_1, X_2) \geq (Y_1, Y_2)$

, or $(Y_1, Y_2) \geq (X_1, X_2)$, or both, in which case the consumer is indifferent between the two bundles.

Reflexive: We assume that any bundle is at least as good as itself: $(X_1, X_2) \geq (X_1, X_2)$.

Transitive: If $(X_1, X_2) \geq (Y_1, Y_2)$ and $(Y_1, Y_2) \geq (Z_1, Z_2)$ then we assume that $(X_1, X_2) \geq (Z_1, Z_2)$. In other words, if the consumer thinks that X is at least as good as Y and that Y is at least as good as Z, then the consumer thinks that X is at least as good as Z.

The first axiom, completeness, is hardly objectionable, at least for the kinds of choices economists generally examine. To say that any two bundles can be compared is simply to say that the consumer is able to make a choice between any two given bundles. One might imagine extreme situations involving life or death choices where ranking the alternatives might be difficult, or even impossible, but these choices are, for the most part, outside the domain of economic analysis.

The second axiom, reflexivity, is trivial. Any bundle is certainly at least as good as an identical bundle. Parents of small children may occasionally observe behavior that violates this assumption, but it seems plausible for most adult behavior.

The third axiom, transitivity, is more problematic. It isn't clear that transitivity of preferences is necessarily a property that preferences would have to have. The assumption that preferences are transitive doesn't seem compelling on grounds of pure logic alone. In fact it's not. Transitivity is a hypothesis about people's choice behavior, not a statement of pure logic.

3.2 Utility

The term utility is unfortunate not just because it suggests usefulness, but because it makes the economic approach to behavior appear more limited than it actually is. We will make very few assumptions about the form of utility that a consumer might have. That is, we will attempt to avoid making value judgments about the preferences a consumer holds – whether they like smoking cigarettes or eating only carrots, watching Arnold Schwarzenegger movies or spending time with a hula hoop. Consumers like whatever it is that they like; the economic assumption is that they attempt to obtain the goods that they like. It is the consequences of the pursuit of happiness that comprise the core of consumer theory. Thus, before discussing the concept of utility, let us first point out some of the assumptions that economists make about the average consumer.

- Economists assume that an average consumer is rational .i.e.
 - A consumer has a clear-cut preference. By this we mean that the consumer is able to compare any two bundles 'a' and 'b' and decide which one he/she prefers
 - A consumer has a persistent/consistent preference. If consumer prefers bundle 'a' to 'b' and 'b' to 'c', then he/she prefers 'a' to 'c'. i.e. transitivity
 - The consumer is not free in his/her choice. The consumer's choice is limited (constrained) by his/her income level and the prices of goods and services. In short, consumer is subject to his/her budget constraint.

Definition of Utility: it is the satisfaction or pleasure that a consumer can/will derive from consuming different goods and services. Alternatively economists use the term utility in a peculiar and idiosyncratic way. i.e., utility refers not to usefulness but to the flow of pleasure or happiness that a person enjoys – some measure of the satisfaction a person experiences. Usefulness might contribute to utility, but so does style, fashion, or even whimsy.

A rational consumer build-up a demand for a product- or will have willingness and ability to purchase a given product expecting a satisfaction as a reward. And the more goods and services it is having, the higher will be the level of satisfaction. Had not been there a limit of resources /in return of goods and services/, any person might enjoy infinite level of satisfaction. However, goods and services are limited, and hence utility is to be limited. Economists have a science here to maximize the level of utility as much as possible, allowing constraints to work of their own field and economists to economize the limited space - and here is the **Economics of Utility**.

In summary, Utility is the power of the product to satisfy human wants .In strict sense of the word; utility does not mean (realized) satisfaction. It refers to expected satisfaction. Hence, when a consumer decides to buy certain quantities of goods and services, he /she consider the expected satisfaction. But, to make our decision simple, we usually assume that the expected satisfaction would be the same as the realized satisfaction. a simple observation from reality shows that consumers obtain utility from their consumption of goods and services only part by part. Consumers consume goods and services bit by bit in pieces .Therefore; we can assume that utility is obtained as consumers gradually built it up. This is what we called total utility.

Definition of Total utility (TU): it is the total accumulated level of satisfaction obtained after a consumer consumed a certain goods and services in a given period.

Definition of Marginal utility (MU): it refers to the level of utility obtained from consumption of each additional consumption of goods and services.

Once we understand the concept of utility is all about, the next question is how to do consumers measure the satisfaction level they derive from goods and services .Is utility measurable? Given his /her budget constraint, how can a consumer buy the different quantities of goods and services that can give him/her the maximum possible utility? This will be answered by the discussion of the theory of utility in the following subsection.

3.2.1 Approaches to the Economics of Utility (Theory of Utility)

We have an agreement to treat utility based on economic principles- and it is actually done by former economists though they are on two broad categories. In this section, we will see the two theories of utility – the Cardinal and ordinal theory of utility.i.e. we will try to see the two different approaches developed to measure or compare consumer's utility derived from the consumption of goods and services. To see how each of the above theories explains the consumer's behavior, let us examine each of them separately.

1. The Cardinal Utility Theory

According to the cardinal utility theory, utility is measurable by arbitrary unit of measurement called **utils** in the form of 1, 2, 3 etc. Utils- is a hypothetical measuring unit of utility forwarded by cardinalists.

Cardinalists- are those that treat utility to be measurable and quantifiable like any other physical object- 10 kg of sugar etc. For example, consumption of an orange gives Bilen 10 utils and, a banana gives her 8 utils, and so on. From this, we can say that an orange gives Bilen more utility than that of banana. Hence, it is not only possible to compare a satisfaction of different goods and services like orange is greater in utility than banana, but also showing by how much quantity actually it did- orange is greater by 2 utils than banana. So, according to cardinal utility theory, utility is measurable or comparable in number /quantities. But, this theory is true if we state the following assumptions.

Assumptions of Cardinalists

- **Independence of utility:** Utility from consuming a certain item is neutral from the satisfaction to be gained from another product by the same consumer. Total level of satisfaction is therefore purely an additive.
- **Constancy of marginal utility of money:** A given unit of money deserves the same value at any time or place it is to be spent. A person at the start of the month where he has received monthly salary gives equal value to 1 birr with what he may give it after three weeks or so.

As stated above, utility is the satisfaction or pleasure that a consumer obtains from consuming a product. Total utility (TU) is the total satisfaction or pleasure a consumer derives from consuming a specific quantities of a commodities at a particular time. And marginal utility is the extra satisfaction that a consumer obtains from consuming additional good and services.

Mathematically, marginal utility (MU) = $(\Delta \text{ TU})/(\Delta Q)$,

Where, $\Delta \text{ TU}$ represents change in total utility, and ΔQ refers to change in the amount of the product consumed. Let's illustrate total utility and marginal using table and graph form.

Table 1: Total and Marginal Utility

Banana consumed per hour	Total utility in utils (TU)	Marginal Utility (MU)

0	0	-
1	10	10
2	18	8
3	24	6
4	28	4
5	30	2
6	30	0
7	28	-2

Relationship between TU and MU

Cardinalists believe that utility to be additive. Whenever a person consumes, there is an independent utility for that particular product which is technically called **Marginal Utility**. The whole of MU will compose the total satisfaction technically called as **Total Utility**.

- *The total utility first increases, reaches the maximum (when the consumer consumes 6 units) and then declines as the quantity consumed increases.*
- *On the other hand, the marginal utility continuously declines (even become zero or negative) as quantity consumed increases.*

If the marginal utility is to be positive, total utility will increase by some amount. Then, we can generalize:

- If MU is positive, TU will increase
- If MU is negative, TU will decrease
- If MU is zero , TU is at maximum

Graphically, the above data can be depicted as follows

Level of Utility

30

U

18

0 2 6 Quantity Consumed

Level of Utility

8

0 Quantity Consumed

Fig 1: Total and marginal utility curves

In the same manner from the above graph, we can see that TU was increasing up to the **6th** level of output where MU is positive. TU reaches its maximum at the 6th unit and then decreases whenever MU is negative where TU starts to fall.

The other point to be remembered is TU may assume different shapes when MU is positive but decreasing and increasing still being a positive number. The conditions:

1. When MU is increasing and positive (actually MU can increase only when it is positive), TU will not only increase but also at an increasing rate. In this region, TU will open up wards.
2. When MU decreases at its positive region, TU will increase at a decreasing rate.

The Law of Diminishing Marginal Utility (LDMU)

This law states that , other things remaining constant , consuming successive units of a product gives a consumer less and less extra satisfaction .In other words, the extra satisfaction(MU) that a consumer derives declines as he/she consumes more and more of the product in a given period of time. Here, the more units of a commodity the individual consumes per unit of time, the greater is the total utility he receives. Although total utility increases, the extra, or marginal, utility received from consuming each additional unit of the commodity decreases. It is also defined as the additional benefit which a person derives from a given increase of his stock of a thing diminishes with every increase in the stock of what he already has. This gives sense in that the first banana a person consumes gives him more marginal utility/MU/ than that of the second and the second banana also gives him higher marginal utility than that of the third and so on.

Why?

While a person may have an unlimited want, each single want is satiable.

Briefing the question

A question here may be, what a contradictory principle it is to say unlimited want in one hand and law of diminishing marginal utility on the other?

Do you feel the same contradiction?

Whether you feel or not there is no red light to indicate danger as you can explore right down here starting from the following short story of ***Mr Johnny Walker***.

There was a person called Johnny 'Walker' who works in a whiskey factory. For some back ground history, he was allowed to drink whatever bottles of whiskey he wants.

And he did so for many years till his death drinking many bottles of whiskey- but which is finite and limited up to his maxima. Can we say that Johnny walker was a person whose all wants are satisfied being exceptional from anybody else in the world? ***Really not the case!***

You see want of Johnny Walker is saturated only by his drinking of a special brand of whiskey.

What if he wants to drink other brands of whisky? What about his desire of Cars, House etc? So still having unlimited wants but regarding drink of that brand of whiskey he is saturated enough.

Therefore, we can generalize, as the person is having more and more of a given product, his intensity of want for that product will diminish, and after some level it may reach zero and even negative, like what is vomit for too much drinking

When an individual consumes more and more units of a good, the intensity of his particular want for the good diminishes. For example, a person has consumed 8 units of banana and ready for the ninth one but no bread at all. The last penny spent for banana might not provide him as equal pleasure than if he spend it for the first loaf of bread. Why? Some level of banana is saturated better than the very empty jar of bread. Now, the feeling of contradiction before is already dead and we are wholly cognizant of it... Yes that is the beauty of reading. ***Go on reader!***

But this law is based on the following assumptions

- **The consumer is rational**
- **The consumer consumes identical or homogenous product**
- **There is no time gap in the consumption of two goods ,and**
- **The consumer taste / preferences remain unchanged .Consuming successive unit of a product over a period gives less and less marginal utility.**
- **The commodity to be consumed should have similar quality, color, design etc.**

When will MU violate the law of diminishing utility marginal (LDMU) and increase while output is increasing?

You missed your breakfast in the morning at the campus cafeteria. Sad to stay at introduction to economics class ended up at 12:30 AM. As you were released from class, you were hurrying in to the cafeteria. Can only one swallow of “*injera*” with ‘therefore-’ erode away your hunger? Frankly and definitely, it will not except to make you salivate more and more. However, the next 2nd, 3rd, and 4th swallows will do at least some. Therefore, the MU of 2nd unit might be greater than the first and so if each unit is very small to make up a significant utility. If you catch up the drift, here is one of the infrequent possibilities for an exception to LDMU. Initially up to some level of output, MU may increase and TU increases at an increasing rate. The shape of TU being opened upwards shows us this fact of increasing at an increasing rate. However, after certain level of output, MU may decrease though positive. As a result of this fact, TU keeps on increasing though at a decreasing rate. At point of zero marginal utility, TU reaches its maximum, and when MU becomes negative, TU continues to fall.

However, the utility that two individuals derive from consuming the same level of a product may not be the same .Thus; we can say utility is subjective or unique to individuals.

The basic economic model of consumer behavior is that people choose the best things they can afford. But, what is ‘best’?

Equilibrium Point under Cardinal Utility Approach (Equi-Marginal Utility)

The objective of a rational consumer is to maximize total utility .As long as the additional unit consumed brings a positive marginal utility, the consumer wants to consume more of the product because total utility increases .But, given his income limit, and the price level of goods and services .what combination of goods and services should he consume so as to get the maximum total utility.

The principle of Consumer’s equilibrium stated that a consumer gets the maximum total utility good if he allocates his/her income in such a way that the last Birr spent on each yield equal marginal utility.

The equilibrium condition of a consumer that consumes a single good X occurs when the marginal utility of X is equal to its market price.

Proof

The utility function is:

If the consumer buys commodity X, then his expenditure will be. Then, the consumer maximizes the difference between his utility and expenditure

The necessary condition for maximization is equating the derivative of a function to zero. Thus,

In general, for the case of two or more goods, the consumer equilibrium is achieved when the marginal utilities per money spent are equal on each good purchased and his money income available for the purchase of the goods has been exhausted. That is,

Mathematically, the condition of maximization of satisfaction is

A. $MUX/PX = MUY/PY = MUZ/PZ$ and

B. His or her income allocated for consumption is exchanged, i.e.

$$P_X Q_X + P_Y Q_Y + \dots + P_Z Q_Z = M$$

Where: MU_x , MU_y and MU_z are marginal utilities of the three goods X, Y and Z

P_x , P_y and P_z are prices of X, Y and Z

Q_x , Q_y and Q_z are quantities of goods X, Y and Z purchased

M represents income allocated for consumption

The limited income of a person is to be allotted among many goods and services. The process of adjustment is to take place until the above condition is realized. From the assumption of constancy of utility of money, each unit of money is equally desirable to the consumer. Whether it is the last penny for the 8th banana or for the 1st loaf of bread, it is the same for the person in regard to money- but might not be for the utility.

The marginal utility of a good for each unit of money can be given by the ratio MUX/PX . If $MUB_a/pba < MUB_r/pbr$ (where, 'ba' - referring to banana and 'br' - referring to bread), the person will get net benefit in regard to satisfaction by allotting some amount of money away from banana to the loaf of bread. This adjustment cannot be meaningful only when the ratio for the different products is equal- and where cardinalists are saying it to be equilibrium point of a consumer, i.e., $MUB_a/pba = MUB_r/pbr$ (**consumer's equilibrium**)

Example: consider that Bilen has Birr 10 and she consumes two goods, X and Y. The price of good X (P_x) is Birr 1 and the price of Y (P_y) is Birr 2. The total utility (TU) she gets from consumption of the goods are given below.

Table 2: Comparison of Total Utility and Marginal Utility per Birr

Good X				Good Y			
	P_X				P		
Birr 1				$Y = \text{Birr } 2$			
Q	X TU	X MU	X/PMU	Quantity	Y TU	Y MU	MU
u			X				
a							
n							
t							
i							
t							
y							

0	0	-	-	0	0	-	-
1	10	10	10	1	24	24	12
2	18	8	8	2	44	20	10
3	25	7	7	3	62	18	9
4	31	6	6	4	78	16	8
5	36	5	5	5	90	12	6
6	40	4	4	6	96	6	3
7	43	3	3	7	100	4	2

Bilen spends her income of Birr 10 on the two goods, X and Y. To spend her income, Bilen has to compare the MU X per Birr, (MUX/PX) with that of good Y, (MUY/PY). She first buys that good which gives her higher MU per Birr. For example. it would be good for Bilen to spend on the first unit of good Y than good X because MU Y per Birr is higher for the first unit of good Y than that of X(i.e. $12 > 10$).

The following table summarizes her decision making process in identifying the utility maximizing combination of goods.

Table 3: Consumer's decision process

Purchase	Available choice	Decision	Remaining income
For the first purchase	First unit of Y,i.e. $MUY/PY = 12$	Buy first unit of Y	$10 - 2 = 8$
For second purchase	Second unit of Y and first unit of X, ($MUX/PX=MUY/PY=10$)	Buy second unit of Y and first unit of X	$8 - (2+1) = 5$
For the third purchase	third unit of Y, $MUY/PY = 9$	Buy third unit of Y	$5 - 2 = 3$
For the fourth purchase	Fourth unit of Y and second unit of X, ($MUY/PY=MUX/PX=8$)	Buy fourth unit of Y and second unit of X.	$3 - (2+1) = 0$

Given her income level, Bilen gets the maximum total utility when she buys 2 units of X and 4 units of Y .Why? Because the last Birr spent on each good gives her equal marginal utility when she buys these levels of the two goods, and at that point, her income is exhausted.

Numerical example,

$$C. MUX/PX = MUY/PY (8/1=16/2=8$$

$$D. PX.QX+PY.QY= M((1*2)+(2*4)=10(2+8= 10(10=10$$

The total utility of Bilen derives from this combination would be the maximum, i.e.

$$TU= TUX + TUY$$

$$96 = 18 + 78$$

Given her fixed income, the price level of the two goods, no combination of the two goods will give her higher TU than this level of utility.

Limitations of cardinal Utility Analysis

LXXX. Non constancy of marginal utility- can a person treat a unit of money at times of having 100 birr and 1 birr equally.

LXXXI. When price is decreasing, real income of a person will increase and each unit of money will deserve high, no more constancy.

LXXXII. Utility is non-measurable.

LXXXIII. Utility is not wholly independent

Could you have the same utility for a cup of coffee at exam and vacation times?

2. The Ordinal Utility Theory or Indifference Curve Approach

Ordinalists- highly suspect for the possibility of utility to be measurable as it is largely of psychic nature. They believed rather it is possible and correct to rank rather than quantifying. So, unlike the cardinal theory, the ordinal utility theory says that utility is not measurable. Rather the consumer can rank or order the utility he derives from different goods. The ordinal utility theory states that utility is not measurable in the form of 1, 2, 3 .etc..According to this theory, it is meaningless to attach numbers to utility we derive from a product .Rather, a consumer can order or rank goods he/she would consume in the form of 1st, 2nd, 3rd etc. This theory provides another method of studying the consumer's behavior .Since it uses indifference curves to study the consumer's behavior, the ordinal utility theory is known as the indifference curve approach.

Indifference Curve

An indifference curve is a locus of points in commodity space- or commodity bundles- among which the consumer is indifferent each point on an indifference curve yields the same total utility as any other point on the same indifference curve. If the utility function is given by:

$$U = (X_1, X_2, X_3, \dots, X_n)$$

where X_1 is the amount of good-1 consumed, X_2 the amount of good- consumed, and so on, then an indifference curve is defined as a set of all commodity bundles $(X_1, X_2, X_3, \dots, X_n)$ that satisfy the equation $U(X_1, X_2, X_3, \dots, X_n) = C$, where C is the constant level of utility for that indifference curve.

The indifference curves approach considers the following assumptions:

84. Consumer's objective is to maximize his/her utility. i.e. the consumer is rational.
2. Consumers can simply order their preferences or they only rank the utility level that they derives from different consumption bundles, i.e. they are not involved in quantifying the different level of utility they obtain from various combination of goods.
3. The rate at which one good can be substituted for another in consumer's basket of goods diminishes as the consumer consumes more and more of the goods ,i.e. there is diminishing marginal rate of substitution in the consumer's consumption bundles
4. The total utility of a consumer is measured by the amount / quantities of all items he/she consumed from his/her consumption basket
5. Consumer's preference for items in his/her consumption bundles is consistent .For example , if there are three goods in a given consumer's basket say , X,Y,Z and if he prefers X to Y and Y to Z , then the consumer is expected to prefer X to Z .this property is known as axioms of transitivity .

Indifference Set

A combination of goods for which the consumer is indifferent is called indifference set. It shows the various combinations of goods from which the consumer derives the same level of utility. Consider a consumer who consumes two goods.

Table 4: Indifference set

Combinations	Quantity of Good X	Quantity of Good Y
A	10	2
B	6	4
C	3	6
D	2	8

Each combination of good X and Y gives the consumer equal level of total utility .Thus, the individual is indifferent weather he consumes combination 'A' , 'B' , 'C' or 'D'.

When the indifference set is expressed graphically, it called an indifference curve. It shows all the combination of the two goods, which gives the consumer equal total utility. Thus, **an indifference curve contains equally preferred bundles. Equal preference (same utility level. Therefore, all bundles in an indifference curve have the same utility level.**

Quantity of good Y



Fig2: Indifference curve for two products X and Y

At point A, a consumer can consume 5 units of good Y and 10 units of good X, at point (C), 3 units of good Y and 25 units of good X , at point (D) 1 unit of good Y and 70 units of good X. At all points the consumer obtains equal level of total satisfaction if even the consumer consumes different amount of goods at each point.

Indifference map: it is a set of indifference curves. Each successive curve to the right represents higher level of total utility

Fig3: Indifference map

Characteristics (properties) of Indifference Curves

- **Slopes downward to the right:** this implies that when the consumer given more of one bundle, she/he has to be given less of the other bundle to leave the consumer on the same level of satisfaction.

- **Convex to the origin** - because of the law of diminishing marginal utility, the more we consume a product, the lower will be its marginal utility. This product should be taken at a higher amount so as to compensate for any lost satisfaction. Alternatively, Indifference curves have a convex shape which is to mean less of a product found to be at a less unit is costly in terms of the other good. A consumer requires a relatively higher amount of another good to release a given product (let X) when it is a smaller level of consumption than when it was a large unit of consumption. The convexity can also be given by the fact of increasing marginal rate of substitution. The marginal rate of substitution says the ratio of the amount to be forgone to the amount to be consumed will fall as we are having less of what is to be forgone. If we are left with a very minimal level of what is to be forgone, it will be hard for us to substitute it with consumption of the other where already have large consumption of it unless it is with a very large proportion.

- **The higher an Indifference curve is the higher will be the level of satisfaction-** any indifference curve which is to the right is composed of more of goods and services. And any rational consumer will be interested to consume more if he is granted with higher level of satisfaction- so that we can say the higher an indifference curve is the higher will be the level of satisfaction.

- **Indifference curves never cross with each other (cannot intersect)** - the properties of consistency and transitive will limit us from having indifference curves which will cross with each other.

Good Y

Good X

Fig4: crossed indifference curves

From the above graph, the consumer preference B bundle compared to bundle C. On the other hand, following indifference curve 1 (IC1), the consumer is indifferent between bundle A and C, and along indifference curve 2 (IC2) the consumer is indifferent between bundle A and B. According to the principle of transitivity, these implies the consumer is indifferent between bundle B and C which is contradictory or inconsistent with the initial statement where the consumer prefer bundle B to C. Therefore, indifference curves never cross each other.

3.3. Marginal Rate of Substitution

It is a rate at which consumers are willing to substitute one commodity for another in such a way that the consumer remains on the same indifference curve. It shows a consumer's willingness to substitute one good for another while the individual is indifferent between the bundles. To understand the concept, take the examples, represented by following figure.

Good Y

Good X

Fig5: Indifference curve for two products X and Y

For an indifference curve labeled IC, different combinations leave the consumer indifferent. When the consumer shifts from point A to point B combination, there is a tradeoff between the consumption different units of bundles.

Mathematically,

$$MRS = \Delta Y / \Delta X$$

This shows the amount of good Y scarified in order to obtain a given amount of good X. Since one of the good is scarified to obtain more of the other good, the MRS is negative. However, we take the absolute value of the result. i.e.

$$MRS_{x,y} =$$

From the above graph, the marginal rate of substitution of moving from point A to B is $\Delta Y / \Delta X = 10/5 = 2$. On the other hand, MRS of moving from point C to D is $4/5$. That is, for the same increment of consumption of good X, the amount of good Y scarified decreases as the consumer consumes more of good X. As a consumer consumes more and more of good X, the amount of good Y the consumer is willing to scarify diminishes. This principle of marginal rate of substitution is reflected by the slope of the indifference curve and is called diminishing marginal rate of substitution.

3.1. Budget Constraint

We begin by examining the concept of the **budget constraint**. Suppose that there is some set of goods from which the consumer can choose. In real life there are many goods to consume, but for our purposes it is convenient to consider only the case of two goods, since we can then depict the consumer's choice behavior graphically.

Given income (M) of the consumer and price of the two products (Px, Py), budget constraint of the consumer requires that the amount of money spent on the two goods be no more than the total amount the consumer has to spend. The consumer's affordable consumption bundles are those that don't cost any more than M. We call this set of affordable consumption bundles the **budget set** of the consumer.

$$P_x X + P_y Y \leq M$$

Budget line - it is a set of the commodity bundles that can be purchased if the entire money income is spent. It is a graph that shows the various combinations of two goods that a consumer can buy given his/her limited income and the prices of the goods. It is given as

$$P_x X + P_y Y = M$$

P_xX is consumer's expenditure on good X, P_yY is consumer's expenditure on good X. If the consumer spends all of his/her income (M) on good X, (Y=0 so that P_yY=0), then P_xX + 0 = M (P_xX=M). Therefore, X=M/P_x (the horizontal intercept. On the contrary, if the consumer spends all of his/her income (M) on good Y, (X=0 so that P_xX=0), then P_yY + 0 = M (P_yY=M). Therefore, Y=M/P_y (the vertical intercept).

When we put good X on the horizontal axis and good Y on vertical axis and joint the two intercepts by a straight line, we obtain the budget line. We can say that,

- Any combination of the two goods on or within the budget line are attainable
- Any combination of two goods above the budget line is unattainable to the consumer, i.e. he/she cannot afford to buy them.

Graphically

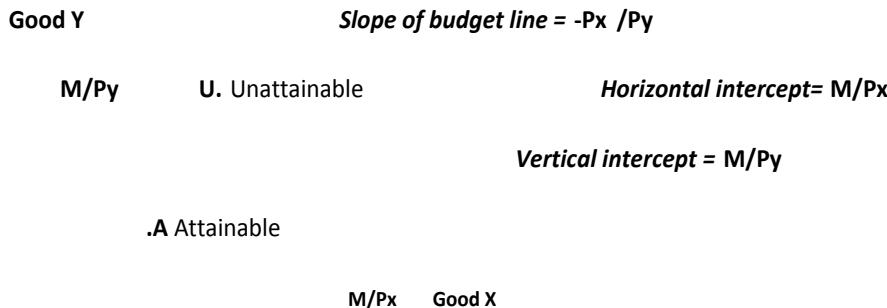


Fig6: Budget line

Given the budget line $P_x X + P_y Y = M$, this can be reduced to the form $Y = (M/P_y) - (P_x/P_y) X$. Then taking the change in vertical (ΔY) over change in the horizontal (ΔX) will give the slope of the budget line ($-P_x/P_y$) which is the ratio of prices of the two goods. The slope shows the market's willingness [not the consumer's willingness] to substitute one good for another.

Factors that influence the budget line to change

The budget line can be shifted by allowing changes on one of determinants of budget line like changes on consumer's income and price of the goods. Taxes, subsidies, and rationing also change the slope and position of the budget line by changing the prices paid by the consumer. Here, if one of the price or both of the goods' price change(s), we only expect and observe change on the slope of the budget line. Let's look at each of determinant influence on our budget line.

Case 1: Income changes

An increase in the consumer's income shifts the budget line outward (to the right). Because, the purchasing ability of the consumer increases. On the other hand, a decrease in income shifts the budget line inward.

Fig 7: Effects of increase (right) and decrease (left) in income on the budget line

Case 2: Price changes

An equal increase in the prices of the two goods shifts the budget line inward. Since the two goods become expensive, the consumer can purchase the lesser amount of the two goods. An equal decrease in the prices of the two goods, on the other hand, shifts the budget line outward. Since the two goods become cheaper, the consumer can purchase the more amounts of the two goods

Fig 8: Effect of proportionate increases (inward) and decrease (outward) in the prices of both goods.

An increase or decrease in the price of one of the two goods, keeping the price of the other good and income constant, changes the slope of the budget line by affecting only the intercept of the commodity that records the change in the price. For instance, if the price of good X decreases while both the price of good Y and consumer's income remain unchanged, the horizontal intercept moves outward and makes the budget line flatter. The reverse is true if the price of good X increases. On the other way, if the price of good Y decreases while both the price of good X and consumer's income remain unchanged, the vertical intercept moves upward and makes the budget line steeper. The reverse is true for an increase in the price of good Y.

Good Y

Good X

Fig 9: Effect of decrease in the price of only good X on the budget line

Note: income and prices might change at the same time affecting the budget line.

Case 3: Taxes, Subsidies, and Rationing

Economic policies (taxes, subsidies, rationing) often are tools that affect a consumer's budget constraint. For example, if the government imposes a quantity tax, this means that the consumer has to pay a certain amount to the government for each unit of the good he purchases.

- **Quantity (volume) tax** – taxes imposed per physical unit of the good purchased. If t birr is imposed on each unit of commodity X purchased, the consumer pays an amount of money equal to the price of commodity X and the tax. That is, the price increases from P_x to $(P_x + t)$. From the viewpoint of the consumer the tax is just like a higher price.
- **An ad valorem (value) tax** – tax imposed on the value of a good rather than on the quantity purchased. A value tax is usually expressed in percentage terms. If a tax rate of t percent is imposed on good X, the consumer's has to pay has to pay P_x to the supplier and tP_x to the government for each unit of the good so the total cost of the good to the consumer is $(1 + r)P_x$. Like the quantity tax, the effect of value tax is just like a higher price.

A subsidy is the opposite of a tax.

- **Quantity subsidy** - the government gives an amount to the consumer that depends on the amount of the good purchased. If, for example, the consumption of wheat were subsidized, the government would pay some amount of money to each consumer of wheat depending on the amount that consumer purchased. If the subsidy is s birr per unit of consumption of good X, then from the viewpoint of the consumer, the price of good X would be $P_x - s$. This would therefore make the budget line flatter.

- Similarly, **an ad -valorem subsidy** is a subsidy based on the price of the good being subsidized. If the government provides a subsidy at a rate of s percent on good X, then the actual price of good X facing the consumer will be $(1 - s)P_x$.

You can see that taxes and subsidies affect prices in exactly the same way except for the algebraic sign: a tax increases the price to the consumer, and a subsidy decreases it. Another kind of tax or subsidy that the government might use is a lump sum tax or subsidy. In the case of a tax, this means that the government takes away some fixed amount of money, regardless of the individual's behavior. Thus a lump-sum tax means that the budget line of a consumer will shift inward because his money income has been reduced. Similarly, a lump-sum subsidy means that the budget line will shift outward.

Governments also sometimes impose **rationing** constraints. This means that the level of consumption of some good is fixed to be no larger than some amount. Suppose, for example, that good X were rationed so that no more than X_1 could be consumed by a given consumer. In this case the budget line will be lopped off at the point.

Good Y

Good X

Fig 10: Effect of rationing on the budget line

3.5 Utility Maximization (Consumer's Equilibrium using indifference curve)

Consumer choice

Given consumer preference and budget constraints, we can now determine how individual consumers choose how much of each good to buy we assume that consumers make this choice in a rational way- that they choose goods to maximize the satisfaction they can achieve, given the limited budget amount available to them.

From the above discussion, we can understand that the indifference curves indicate the desires of the consumer. It shows various combinations of goods that give the consumer equal level of total satisfaction. On the other hand, the budget line shows that the feasible or attainable combinations of the goods at certain level of income and prices. Now, let us bring the indifference curves and the budget line together so as to

identify that combination of goods which will give the consumer the maximum possible satisfaction .This combination is known as satisfaction unit or consumer's equilibrium point.

A consumer maximizes the total utility or satisfaction obtained from spending his income and is in equilibrium when the marginal utility of the last dollar spent on each commodity is the same. This equilibrium condition for utility maximization can be restated as follows:

Quantity of good Y

20 **Budget line**

Indifference curve

E

IC

30

Quantity of good X

Fig11: Consumer's equilibrium for single indifference curve

In general, consumer obtains the maximum possible total utility when he buys that combination of the goods at which the budget line is tangent to the highest attainable indifference curve.

Good Y

Highest possible attainable indifference Curve

20 c
a b

IC3

IC1 IC2

30

Good X

Fig12: Consumer's equilibrium for various indifference curves

Mathematically , at point 'b' above, the slope of budget line = the highest attainable indifference curve

The slope of budget line is the ratio of prices, i.e. P_y/P_x

The slope of the indifference curve is **MRS X for Y**, i.e.

Thus, it is true at point' b', $P_y/P_x = MRS X for Y = (consumer's equilibrium point)$

However, combination at point 'a' is attainable in terms of resource but provides consumer relatively low level of total satisfaction because it is IC1 which is lower than IC2. Combination at 'c' gives consumer relatively more level of total satisfaction because it is IC3 which is higher than both IC1&IC2. But, it is unattainable since it lies outside the budget line.

From the previous section, we have identified that there will be a specific point for consumer to maximize utility. As income is changing, the budget will also shift to either of the directions and also the point of equilibrium for the consumer.

Review Questions

99. Elaborate the justifications for the downward sloping and convexity of indifference curve.
2. Suppose that the set of indifference curve is not downward sloping, what can we say about the desirability of the two goods?
3. Draw a set of indifference curves for which the marginal rate of substitution is constant. Draw two budget lines with different slopes; show what the satisfaction-maximizing choice will be in each case. What conclusions can you say?
4. The price of good X is \$ 8 and the price of good Y is \$ 10. Kuntakunti has a budget of \$ 80 and has already bought 4 good X. He has remained still with \$ 48 more to spend on additional good X and Y. Draw this budget line. If his consumption expenditure made on 1 good X and 4 good Y. Show Kuntakunti's consumption choice on the budget line.
5. The law of diminishing marginal utility states that each successive unit of the commodity consumed leads to a smaller addition to total utility. Why?
6. Originally, the consumer faces the budget line $P_1 X_1 + P_2 X_2 = M$. Then, the price of good 1 doubles, the price of good 2 becomes 8 times larger, and income becomes 4 times larger. Write down an equation for the new budget line in terms of the original prices and income.
7. What happens to the budget line if the price of good X increases, but the price of good Y and income remain constant? Show graphically
8. If the price of good 1 doubles and the price of good 2 triples, does the budget line become flatter or steeper? Show graphically
9. Suppose that the government puts a tax of 15 cents a gallon on gasoline and then later decides to put a subsidy on gasoline at a rate of 7 cents a gallon. What net tax is this combination equivalent to?
10. Suppose that a budget equation is given by $P_1 X_1 + P_2 X_2 = M$. The government decides to impose a lump-sum tax of u , a quantity tax on good 1 of t , and a quantity subsidy on good 2 of s . What is the formula for the new budget line?
11. If the income of the consumer increases and one of the prices decreases at the same time, will the consumer necessarily be at least as well-off?
12. Suppose that a consumer always consumes 2 spoons of sugar with each cup of coffee. If the price of sugar is P_1 per spoonful and the price of coffee is P_2 per cup and the consumer has M dollars to spend on coffee and sugar, how much will he or she want to purchase?
13. Suppose Eleni and Haweni have decided to allocate Birr 200 on refreshment in the form of Coffee and Tea per month. But Eleni and Haweni differ in their preferences for the product. Eleni prefer Tea to Coffee, while Haweni prefer coffee to Tea.
 - a. Draw a set of indifference curves for Eleni and a second set for Haweni
 - b. Discuss why the two indifference curves are different from each other using the concept of marginal rate of substitution.
 - c. If both Eleni and Haweni pay the same price for refreshments, will their marginal rate of substitution of coffee for Tea be same or different? Explain.
14. Tamirat buys 2 new college text books during his first year at a school at a cost of Birr 50 each. Used text books cost only Birr 30. When the book store announces that there will be a 20% and 10% increase in price of new text books and used text books respectively for the coming year. As a result, Tamirat's family gives him Birr 80 extra. Is Tamirat better off or worse off after price change?
15. The utility that Florence obtains by consuming food (F) and cloth (C) is given by
$$U=f(F, C)=FC$$
 - a. Suppose that food costs Birr 1 per unit, and clothing costs Birr 3 per unit, and Florence has Birr 12 to spend on food and cloth. Graph the budget line she faces.
 - b. What is utility maximizing choice of food and clothing? (solve the problem both graphically and mathematically)
 - c. What is marginal rate substitution of food for cloth when utility is maximized?

CHAPTER FOUR: THE THEORY OF PRODUCTION

In the previous chapter, we focused on the demand side of the market – the preferences and the behavior of consumers. Now we turn to the supply side of the market and examine the behavior of producers. Consider some of the problems that a company like General Motors faces regularly. How much assembly-line machinery and how much labor should it use in its new automobile plants? If it wants to increase production, should it hire more workers, or should it also construct new plants? Does it make more sense for one automobile plant to produce different models, or should each model manufactured in a separate plant? These questions apply to not only to business firms, but also to other producers of goods and services, such as governments and nonprofit agencies. In this chapter, therefore, we will see how firms can organize their production efficiently.

4.1. Concepts of Production and Production Function

In simple words, **production** refers to the process of transferring inputs into outputs. An input (factor of production) is a good or service that goes into the production of another good or service. In other words, an input is simply anything which the firm buys for use in its production process. Inputs include labor, land, capital and entrepreneurial talent. The end products of the production process are outputs which could be tangible (goods) or intangible (services).

Simple examples of production:

- A furniture-producing firm combines workers labor time, machineries, his organizational skills and various raw materials like wood, metal, etc to produce sofas for sale to its customers.
- A high school uses teachers, books, educational materials (aids), class rooms, the available technology (like plasma tv), etc to provide educational services to students.

The theory of production explains and formalizes the nature of relationships between factors (inputs) used and output. As you might have noticed, the production process does not necessarily involve physical conversion of raw materials into tangible goods. Besides teachers, lawyers, doctors, social workers, consultants, hair-dressers, etc are all engaged in producing intangible goods.

Production function is a technical relationship between inputs and outputs. It shows the maximum output that can be produced with a fixed amount of inputs and the existing technology. A production function may take the form of an algebraic equation, table or graph.

A general equation for production function could for instance be described as:

(4.1)

where Q is the maximum output produced; and $X_1, X_2, X_3, \dots, X_n$ are different types of inputs.

To illustrate, suppose a wheat-producing firm uses labor (L), capital (K), land (S) and entrepreneurship (E). Other inputs such as seeds, fertilizers, insecticides, and the like may be included in one of these large groups of factors of production. The production function for wheat may then be expressed as:

(4.2)

Note that we must assume that the production of Q tons of wheat is realized in the most efficient way possible. If it, for instance, is possible to produce 20 tons of wheat using a certain combination of L, K, S and E , it is also possible to produce only 19 tons with the same combination. So, the second technology has to be abandoned as it is not efficient (it is wasting resources that could be used for the production of one more ton).

In sum, production functions describe what is technically feasible when the firm operates efficiently: that is, when the firm uses each combination of inputs as effectively as possible.

Fixed Vs Variable Inputs

Normally, firms employ inputs whose amount does not change for some time and others whose quantity varies according to the amount of production (output). One can hence categorize inputs as fixed and variable.

Fixed inputs are those inputs whose quantity can not readily be changed when market conditions indicate that an immediate change in output is required. In fact no input is ever absolutely fixed, but may be fixed during an immediate requirement. For example, if the demand for Beer shoots up suddenly in a week, the brewery factories cannot plant additional machinery over a night to respond to the increased demand. It takes long time to buy new machineries, to plant them and use for production. Thus, the quantity of machinery is fixed for some times such as a week. Buildings, plot of land and machineries are examples of fixed inputs because their quantity cannot be manipulated easily in short time periods.

Variable inputs, on the other hand, are those inputs whose quantity can be changed almost instantaneously in response to desired changes in output. That is, their quantity can easily be diminished when the market demand for the product decreases and vice versa. The best example of variable input includes unskilled labor and raw materials.

The Short Run Vs the Long Run

Depending on the nature of economic adjustment in a firm to changing economic environment, the production period is divided into short-run and long-run.

The **short run** refers to a period of time in which one or more factors of production cannot be changed. In this case the firm can alter its level of output by increasing or decreasing the use of variable inputs. A firm's capital, for example, usually requires time to change – a new factory must be planned and built, machinery and other equipment must be ordered and delivered, all of which can take a given time period.

The **long run**, on the other hand, is the amount of time needed to make all inputs variable. Here the period is long enough to allow changes in the level of all inputs.

Note that the supply of fixed inputs in the short-run is *inelastic* while the supply of variable inputs in the short-run is elastic.

How long should the long-run be? A month? A year? 2-3 years? 10 years? ... What do you think? If you attempt to put figures, that is wrong. Sorry, there is no precise answer for the question. In short, it differs from industry to industry and more specifically from firm to firm. In some industries, such as groceries, short-run may be a few weeks or while in some other industries like electricity and telecommunications, short-run may mean 4 or more years. Similarly, long-run may be 2 or 3 years while in other industries it might be 10 or more years. Therefore, long-run and short-run do not refer to any fixed period of time. There is no hard and fast rule that specifies how short is short-run or how long is long-run.

Based on these classifications and concepts, we can see the short-run and the long-run production functions in the sub-sections that follow.

4.2. The Short-Run Production Function: Production with One Variable Input

The majority of production decisions of firms are made in the short-run in which the quantity of at least one factor of production changes with output. Hence, the short-run production function shows the relationship between the maximum product and the level of the variable input. In more general expression, a short-run production could take the following form, for Q output and X₁ variable input quantities:

(4.3)

Imagine, for example, that you are managing a clothing plant. You have a fixed amount of equipment, but you can hire more labor or less to sew and to run the machines. You have to decide how much labor to hire and how much clothing to produce. To make the decision, you will need to know how the amount of output Q increases (if at all), as the input of labor L increases. Therefore, the short-run production function of cloth could be expressed as:

$$\text{Quantity of Cloth} = f(\text{Labor}) \quad (4.4)$$

Here, the short-run level of cloth produced is supposed to depend on labor, the only variable input. Since other factors are assumed to be fixed in the short-term, we do not include them in the production function. This, however, does not mean that they are not used in the production process.

4.2.1. Total, Average and Marginal Products

The contribution that variable input makes to the production process can be described in terms of the total, average and marginal product.

Total Product (TP)

It refers to the total output (say cloth) produced by a given amount of a variable input (say labor) keeping the quantity of other inputs fixed (say machines). In almost all real world production processes, TP in the short-run follows a certain trend: it initially increases at an increasing rate, then increases at a decreasing rate, reaches a maximum point but eventually falls with a rise in the quantity of the variable input. That is, initially, as we combine more and more units of the variable input with the fixed input output continues to increase. But eventually, increasing the unit of the variable input may not help output increase. Even as we employ more and more unit of the variable input beyond the carrying capacity of a fixed input, output may tend to decline. Thus increasing the variable input can increase the level of output only up to a certain point, beyond which the total product tends to fall as more and more of the variable input is utilized. This tells us what shape a total product curve assumes. The shape of the total product curve is nearly S-shape (see fig 4.1 Panel A)

Average Product (AP)

Average product of an input is the level of output that each unit of input produces, on the average. It tells us the mean contribution of each variable input in the total product. Mathematically, AP is total product divided by the amount of variable input used to produce that product. The average product of labor (APL), for instance, is given by:

(4.5)

Like that of TP, APL first increases, gets its maximum value and eventually falls continuously afterwards.

Marginal Product (MP)

We may, at times, be interested in knowing the extra output brought about by the extra employment of a variable input. In terms of labor, we may ask "how much has the last laborer added to total product?" These issues are explained by the marginal concept.

Marginal product is the extra or additional output obtained when one extra unit of a variable input is entered in production while other factors remain fixed. Simply, marginal product is a change in the amount of total product divided by a change in the amount of variable input used. For instance, marginal product of labor (MPL) is given by:

(
4.6)

The last term in this equation (read as the partial derivative of Q with respect to L) is applied when a continuous production function (i.e. an algebraic equation) is given for output Q .

Thus, MPL measures the slope of the total product curve at a given point. In the short run, the MP of the variable input first increases reaches its maximum and then tends to decrease to the extent of being negative. That is, as we continue to combine more and more of the variable inputs with the fixed input, the marginal product of the variable input increases initially and then declines.

Tabular and Graphical representation of the short run product curves

Table 4.1 contains the total product (TP) and other product types for various amounts of labor units (e.g. number of workers) and a constant amount of capital (e.g. 5 machineries). The values in column 3 of the table show total output. If you carefully have a look at the values of the TP, they initially increase at an increasing rate as labor units rise from 0 to 4. TP still increases but at a decreasing rate when labor units are between 4 and 8. The maximum total output of 119 units is realized by combining 8 units of labor and 5 units of capital at the cheaply available technology. This could be interpreted as the total output maximizing level of labor employment is 8. Finally, the total product begins to decline when the amount of the variable input exceeds 8 units.

Table 4.1 Production with One Variable Input

Amo unt of Lab or, <i>L</i>	Am ou nt of Capital, K	T o ta l P ro d u ct T P (Q)	A v e r a g e P r o d u c t A P (=) T P L	Marginal Product MP($=\Delta TP/\Delta L$)	Stage of Production
0	5	0	-	-	Stage I
1	5	1 0	1 0	10	
2	5	3 2	1 6	22	
3	5	6 3	2 1	31	
4	5	8 4	2 1	21	Stage II
5	5	1 0 0	2 0	16	
6	5	1 1 1	1 8 5	11	
7	5	1 1 9	1 7	8	
8	5	1 1 9	1 4 9	0	

9	5	1 1 7	1 3	-2	Stage III
10	5	1 1 0	1 1	-7	

Now, copy the values of the 1st and 3rd columns of *Table 4.1* on a separate piece of paper. Then, use *equation (4.5)* to compute the corresponding values of APL (for each value of labor unit). Compare your results with those in the 4th column. What trend have you observed for the APL?

We can see that the behavior of AP is almost similar to that of TP. Like that of TP, APL first increases up to 4 labor units, gets its maximum value of 21 output units at the 4th labor unit and eventually falls continuously afterwards.

Now also, copy the values of the 1st and 3rd columns of *Table 4.1* on a separate piece of paper. Then, use *equation (4.6)* to compute the corresponding values of MPL (for each value of labor unit). Compare your results with those in the fifth column.

We observe an almost similar behavior for MPL as APL and TP. Hence, MPL first rises up to the 3rd labor unit, gets its maximum value of 31 output units for the 3rd labor unit and eventually falls afterwards. MPL becomes zero when total product gets its maximum. This means that the last factor of production adds nothing to total output. Further employment of labor beyond the zero MPL point even makes it negative implying that total production is falling from its maximum value. In our illustration, this happens after the employment of 8 units of labor. Workers hired after 8th laborer contribute negatively.

These phenomena's can also be depicted in graph as shown in figure 4.1. below.

Figure 4.1 Total, Average and Marginal Product Curves

As the number of the labor hired increases (capital being fixed), the TP curve first rises, reaches its maximum when 8 amount of labor is employed, beyond which it tends to decline. Assuming that this short run production curve represents a certain cloth manufacturing firm, it implies that 8 numbers of workers are required to efficiently run the machineries. If the numbers of workers fall below 8, the machine is not fully operating, resulting in a fall in TP below 119. On the other hand, increasing the number of workers above 8 will do nothing for the production process because only 8 numbers of workers can efficiently run the machine. Increasing the number of workers above 8, rather results in lower total product because it results in overcrowded and unfavorable working environment.

Marginal product curve increases until 3 amount of laborer is employed and then it tends to fall. The MPL is zero at 8 amount of labor (when the TP is maximal); beyond which its value assumes zero indicating that each additional worker above 8 tends to create over crowded working condition and reduces the total product. Thus, in the short run (where some inputs are fixed), the marginal product of successive units of labor hired increases initially, but not continuously, resulting in the limit to the total production. Geometrically, the MP curve measures the slope of the TP. The slope of the TP curve increases (MP increases) up to 3 amounts of labor units, it decreases from 3 to 8 and it becomes negative beyond 8.

The average product curve increases up to labor unit of 4, beyond which it continuously declines. The AP curve can be measured by the slope of rays originating from the origin to a point on the TP curve. For example, the APL at 4 is the ratio of 84 level of output to 4 labor units. This is identical to the slope of ray a.

Relationship between Average Product and Marginal Product

From the previous discussions and figurative presentations, we can notice specific relationships between APL and MPL.

If we just have a close look at *Table 4.1* and *Figure 4.1*, particularly at APL and MPL, we easily come up with the following:

- When $MPL > APL$, APL keeps on increasing. This is what you observe in the entire area labeled as stage I.
- When $MPL = APL$, APL is already at its maximum. One can locate this at the point where MPL and APL intersect (end of stage I and beginning of stage II).
- When $MPL < APL$, APL keeps on decreasing. You see such a scenario in what we labeled as stages II and III.

A simpler way to understand the relationship between average and marginal product is to think of it in terms of grades. Suppose you took only two courses so far and your average score for the two courses is 90 ($=APL$). If your score for an additional (marginal) course is 93 ($=MPL > APL$), your new average will be 91. But, had your additional (marginal) score been 87 ($=MPL < APL$), your new average would be 89. Thus, your marginal score pushes up or down your average product depending on whether the marginal score is above or below the average respectively. The same relationship holds true for APL and MPL.

As we know, in addition to tables and graphs, equations are important tools in economics to understand relationships.

Numerical Illustration:

Suppose that the short-run production function for cut-flower by a certain Ethiopian firm is given by:

where Q - represents the annual quantity of cut-flower produced.

K - annual capital input; suppose $K=5$.

L - annual labor input.

119. Determine the average product of labor (APL) function.
2. At what level of labor does the total output of cut-flower reach the maximum?
3. What will be the maximum achievable amount of cut-flower production?

Solution:

2. We know that when total product (Q) is maximum, MP will be zero. And .

That is, partially differentiating Q with respect only to L and equating it to zero:

(Q – cut-flower level of output will be the maximum if the firm employs 100 units of labor.)

123. Substituting the optimal values of labor ($L=100$) and capital ($K=5$) into the original production function (Q) gives the maximum level of cut-flower production:

4.2.2. Efficient Region of Production in the Short-Run

We are now not in a position to determine the specific number of the variable input (labor) that the firm should employ because this depends on several other factors than the productivity of labor such as the price of labor, the structure of input and output markets, the demand for output, etc. However, it is possible to determine ranges over which the variable input (labor) be employed.

To do best with this, based on the relationship between TP, MP and AP, economists have defined three stages of production. For visual observation of the discussions that follow, refer back to figure 4.1.

Stage I: This stage of production includes the range of variable input levels at which the average product (APL) continues to increase. Stage I goes from the origin to the point where the APL is maximum, which is the equality of MPL and APL . In terms of *Figure 4.1*, this goes to 4 level of labor employment. At 4, the APL is at its maximum value. A unit increase in labor initially has an impact of increasing output at an increasing rate (up to the 3 unit of labor). That is why this stage of production also called the stage of *increasing marginal returns*.

Two explanations may be given for the presence of increasing marginal returns:

CXXIV. There is plenty of fixed input supply compared to the variable input. Therefore, as more and more units of the variable input are added to the fixed input, the fixed input will be more intensively and effectively exploited. Hence, the efficiency of capital or land will increase in proportion to the additional units of variable input, labor for instance.

CXXV. Another explanation: As more and more units of the variable input is used, there will be division of labor which will result in specialization. Specialization leads to higher productivity.

However, this will not persist indefinitely as average output begins to fall and marginal output continues falling which marks the commencement of the second stage in production.

Stage II: This stage of production covers that range of variable input used at which MPL is less than APL and is positive. In this stage, both the APL and MPL of the variable input (labor) are diminishing but positive. According to *Figure 4.1*, this starts from labor unit of 4 and goes up to 8. In other words, stage II goes from the point where APL is at its maximum ($MPL=APL$) to the point where MPL is zero. Here, as the input increases by one unit, output still increases but at a decreasing rate, i.e., each increment of labor generates a smaller increase in output than the last. This continues until output reaches its maximum at the 8 unit of labor. Due to this, the second stage of production is termed the stage of *diminishing marginal returns*.

The reason for decreasing average and marginal products is due to the scarcity of the fixed factor. Once the optimum capital-labor combination is achieved, employment of additional unit of the variable input will cause the output to increase at a slower rate. As a result, the marginal product diminishes. The additional labor will have less and less of the fixed input to work with. This reasoning has led to the emergence of the law of diminishing marginal product, which will be our focus in the upcoming sub-section.

Any further additional labor unit after 8 will result in a decline in output and such a situation happens at the third stage of production.

Stage III: At this stage, an increase in the variable input results in the decline of the total product. Hence, the total product curve slopes downwards and the marginal product of labor become negative. This happens after labor unit of 8. The stage is also known as the stage of *negative marginal returns* to the variable input.

The cause of negative marginal returns is the fact that the volume of the variable inputs is quite excessive to the fixed input to the extent that they get in each others' way – creating a problem of overcrowding. As a result, the total product declines and results in negative marginal product. The saying "too many cooks spoil the broth" seems to well fit this stage of production.

Out of the three stages of production, we can identify the stage which is most efficient at which a rational producer should produce.

Obviously, a firm should not operate in stage III because in this stage additional units of variable input are contributing negatively to the total product (MP of the variable input is negative) because of crowded working environment i.e., the fixed input is over utilized.

Stage I is also not an efficient region of production though the MP of variable input is positive. The reason is that the variable input (the number of workers) is too small to efficiently run the fixed input; so that the fixed input is underutilized (not efficiently utilized)

Thus, the efficient region of production is stage II. At this stage additional inputs are contributing positively to the total product and MP of successive units of variable input is declining (indicating that the fixed input is being optimally used). Hence, the efficient region of production is over that range of employment of variable input where the marginal product of the variable input is declining but positive.

4.2.3. The Law of Diminishing Marginal Product (LDMP)

The LDMP is the major factor behind the relationship between TP, MP, and AP. It states that as the number of units of the variable input increases, other inputs held constant, the marginal product of the variable input declines after a certain point. The law is sometimes called the law of variable proportions.

For instance, as you add more labor to a garden plot of fixed size, the marginal increase in vegetable may increase. Nonetheless, a point is reached where an increase in the use of the variable input yields progressively less and less additional (marginal) product. Each additional unit has, on average, fewer units of the fixed input with which to work.

4.3. The Long-Run Production Function: Production with Two Variable Inputs

We have completed our analysis of the short-run production function in which the firm uses one variable input (labor) and one fixed input (capital). Now we turn to the long run analysis of production. Remember that long run is a period of time (planning horizon) which is sufficient for the firm to change the quantity of all inputs. For the sake of simplicity, assume that the firm uses two inputs (labor and capital) and both are variable. This can be expressed in equation form as:

The firm can now produce its output in a variety of ways by combining different amounts of labor and capital. With both factors variable, a firm can usually produce a given level of output by using a great deal of labor and very little capital or a great deal of capital and very little labor or moderate amount of both. In this section, we will see how a firm can choose among combinations of labor and capital that generate the same output. To do so, we make the use of isoquants and isocosts. So it is necessary to first see what is meant by isoquants and isocosts and their properties.

Isoquants and Isocosts

For the analysis of a production function with two variables factors in the long-run, we make use of a concept known as isoquants. And adding the concept of isocosts to isoquants, we will be able to analyze the equilibrium situation of a producing firm in the long-run.

126. Isoquant

An isoquant is a curve that shows the different technically efficient combination of the two inputs that can produce the same level of output. It is also called equal-product curve or product indifference curve.

Since an isoquant represents those combinations of two inputs which will capable of producing an equal quantity of output the producer will be indifferent between them. The various combinations of labor and capital could be presented, as usual, in tables called isoquant tables or schedules, graphs called isoquant curves and equations. Due to their simplicity and visual expression, we prefer isoquant curves to others.

The following table presents 5 combinations of labor and capital. Each combination of labor and capital produce the same level of output, say 50 tons of wheat.

Table 4.2 Isoquant Table (Schedule)

Input Combination <i>n</i>	Labor (L)	Capital (K)	Maximum Output
A	1	12	50
B	2	8	50
C	3	5	50
D	4	3	50
E	5	2	50

I hope that plotting the above combination between L and K is a very simple task for you. So, your plot of the data on the XY plane gives you the following iso-product or isoquant curve.

Figure 4.2. Isoquant Curve

Here also, whether the firm produces at point A or B or E, the same amount of output is realized. If the firm produces less or more of that amount, it produces at another point and in another isoquant.

Collection of isoquants is known as an isoquant map, as shown in *Figure 4.3*. The further an isoquant is from the origin, the larger the output level it represents as, for instance, 75 quintals of *teff* is greater than 25 quintals.

Figure 4.3 Isoquant Map

Characteristics (properties) of isoquants

Isoquants have most of the same properties as indifference curves. The biggest difference between them is that output is constant along an isoquant where as indifference curves hold utility constant. Most of the properties of isoquants, results from the word '**efficient**' in its definition.

1. Isoquants slope down ward.

Because isoquants denote efficient combination of inputs that yield the same output, isoquants always have negative slope. Isoquants can never be horizontal, vertical or upward sloping. If for example, isoquants have to assume zero slopes (horizontal line) only one point on the isoquant is efficient. See the following figures.

Fig 4.4

A

B

C

Thus, efficiency requires that isoquants must be negatively sloped. As employment of one factor increases, the employment of the other factor must decrease to produce the same quantity efficiently.

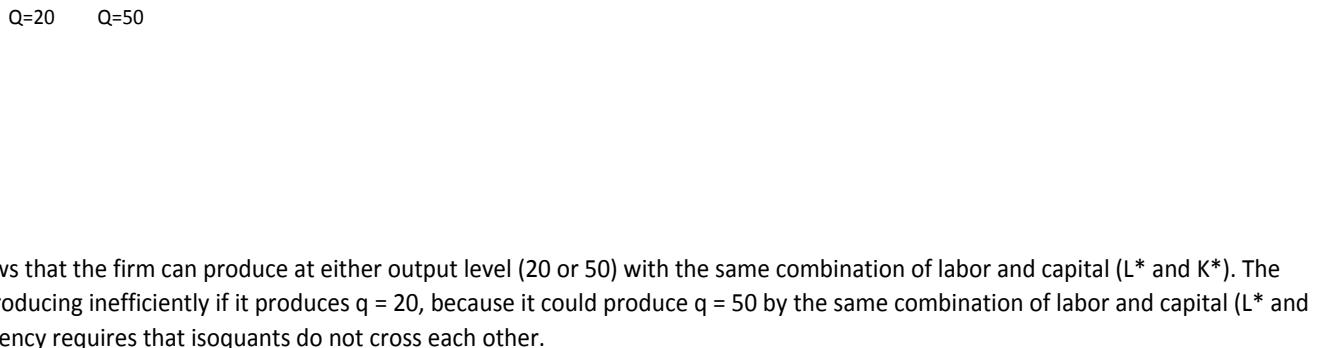
2. The further an isoquant lays away from the origin, the greater the level of output it denotes.

Higher isoquants (isoquants further from the origin) denote higher combination of inputs. The more inputs used, more outputs should be obtained if the firm is producing efficiently. Thus efficiency requires that higher isoquants must denote higher level of output.

3. Isoquants do not cross each other.

This is because such intersections are inconsistent with the definition of isoquants. Consider the following figure.

Q=20 Q=50



This figure shows that the firm can produce at either output level (20 or 50) with the same combination of labor and capital (L^* and K^*). The firm must be producing inefficiently if it produces $q = 20$, because it could produce $q = 50$ by the same combination of labor and capital (L^* and K^*). Thus, efficiency requires that isoquants do not cross each other.

4. Isoquants must be thin.

If isoquants are thick, some points on the isoquant will become inefficient. Consider the following isoquant.

Fig.4.6: *Isoquants can never be thick. Points A and B are on the same isoquant. But point A denotes higher amount of capital and the same amount of labor as point B. Hence point A denotes inefficient combination of inputs and thus it lies out of the isoquant. The isoquant should be thin if point A is to be excluded from the isoquant.*

Marginal Rate of Technical Substitution (MRTS)

It is the rate at which one input can be substituted for another given constant output. It is the absolute value of the slope of the isoquant at the

given point, i.e., marginal rate of technical substitution is the slope of an isoquant.

(3.9)

The d before a variable indicates a change or total differential in the variable.

MRTS and marginal products are highly related. Recall the two-factor long-run production function in *equation (3.8)*.

The total differential of Q is given by:

However, along an isoquant the change in output is zero so that

This means that the ratio of marginal products is also similar to MRTS, i.e.,

(3.10)

Note that the following two things:

A. It is also possible to arrive at the same finding using the following concept. When a producer employs additional labor, he will get additional output given by $MPL(\Delta L)$. When the producer reduces the use of capital input, he will face a decrease in output given by $MPK(\Delta K)$. Since along an isoquant the change in output is zero, the increase and the decrease in output should be equal in magnitude, i.e.,

B. Marginal rate of technical substitution of labor for capital (or) & marginal rate of technical substitution of capital for labor (or) are not identical. Hence,

(3.11)

The Law of Diminishing Marginal Rate of Technical Substitution (LDMRTS)

Most of the time inputs are imperfect substitutes. The LDMRTS states that as the amount of a given input increases, given other inputs, its marginal product decreases. On the other hand when the amount of a given input decreases, its marginal product increases. The more labor the firm has, the harder it is to replace the remaining capital with labor. So, falls as the isoquant becomes flatter.

Table 4.3 Marginal Rate of Technical Substitution

Input Combination	Labor (L)	Capital (K)		
A	1	12	-	0.25
B	2	8	4	0.33

C	3	5	3	0. 5
D	4	3	2	1
E	5	2	1	-

increases as more and more labor is employed – it falls to 1 from zero. It is that we called the law of diminishing marginal rate of technical substitution. On the other hand, increases as the employment of capital increases. This is easily observable from the above table.

Properties of Isoquants

From the isoquant tables, curves and discussions we had so far, we can observe general features or properties about isoquants.

All isoquants share the following common properties.

- C. They are downward sloping (negatively sloped).
- D. They are convex to the origin. This happens because of the law of diminishing marginal rate of technical substitution.
- E. Two isoquants never cross each other. If two isoquants cross each other, a single combination of the two inputs (K and L) will represent two different levels of output, which is false.
- F. Isoquants far from the origin represent higher levels of output.

127. Isocost

As we will discuss in the upcoming chapter, production requires money at least for the acquisition of factors of production. This money is a cost to the producer. A very interesting and simple fact is that a firm can spend an identical total cost for various combinations of inputs. Here comes the concept of an isocost.

An isocost is simply a line that shows the various combinations of two inputs (in our case, labor and capital) that can be purchased for the same amount of outlay (total cost).

Let us assume that the total cost incurred by a firm is only TC . And if the price of labor (L) and price of capital (K) are symbolized by w and r respectively, then

(3.12)

If we express K in terms of others, we get

(3.13)

This equation is known as *isocost equation*. It has the following implications:

- If the firm devotes its entire fund to purchase K , then it can employ TC/r amount of K , leaving no money to hire labor ($L=0$).
- If the firm devotes all the funds to buy labor, then it can hire TC/w amount of L , leaving no money to purchase K ($K=0$).
- All the intermediate positions on these two extreme points show any other combinations of L and K the firm can hire at a cost of exactly TC .

Graphically, this could easily be put as follows.

Figure 4.7 Isocost Line

Slope of an isocost line is given by the ratio of price of labor (w) and the price of capital (r), i.e., it is just the derivative of the isocost equation given by *equation (3.13)* with respect to labor (L).

(3.14)

Note that the slope an isoquant and the slope of an isocost are both negative, implying the use or purchase of one input disregards the other. (We have omitted it deliberately above.)

4.4. Optimum of the Producer in the Long-Run

The objective of any firm is maximizing profit which could be achieved by either output maximization or cost minimization or by both. In general, the optimum of the producer refers to the least cost combination of inputs providing the maximum achievable output level. The optimum levels of the two inputs (K and L) can be found by the use of isoquant curves and isocost lines.

A rational firm wants to produce more output *i.e.* to have isoquants far from the origin (like $Q_3=75$ in *Figure 3.6*). However, that desire may be constrained by shortage of outlay for the purchase of factors of production. The money may not buy enough inputs which help produce no more than 50 units of the output ($Q_2=50$). So, there must be a certain level of employment of inputs which maximizes output at the given cost of production.

The optimum of the producer is achieved at the point where an isocost line is tangent to an isoquant curve. This means the point at which the slope of an isoquant is equal to the slope of an isocost line.

Figure 4.8. Equilibrium of the Producer

- At point A, the firm uses all its money to buy L and K . But, it produces only 25 units of a product; it is very low. The firm's outlay is able to buy more inputs to produce further output. Point B is similar to point A in terms of level of production.
- At point C, the firm can produce 75 units of a product. However, it cannot purchase the combination of L and K that produces this much quantity.
- At point E, the firm produces 50 units of a product; and this is the maximum possible amount given the cost. Therefore, the optimum combination of K and L is found at point E where the isocost line is tangent to the second isoquant curve. At this equilibrium point, the slope of the isoquant equals with the slope of the isocost, *i.e.* .
- Accordingly, the firm has to purchase L^* units of labor and K^* units of capital to realize the maximum possible level of output (=50 units) using the limited outlay of TC .

Numerical Illustration

Suppose a certain small enterprise allocates only $20,000$ *birr* for the production of furniture (school armchairs). The enterprise wants to employ workers (L) whose wage is $w=1000$ *birr* and purchase implements (K) at a price of $r=4000$ *birr*. Suppose further that the production function for furniture is given by .

135. Determine the marginal product functions of workers and implements.
2. Find and .

3. How many workers (L) and implements (K) must be acquired for the small enterprise to produce the maximum possible number of armchairs.
4. How many armchairs will be produced at the equilibrium of the enterprise?

Solution:

We are given the production function as and the total cost (TC) function as .

139. Marginal product of a worker (MPL):

Marginal product of an implement (MPK):

140. and

141. The optimum of the small enterprise is found at the following point:

or .

If we use the first equality, we have:

Substituting this last equation into the TC equation, gives:

Hence,

If the enterprise purchases 4 implements and employs 16 workers, it will have the least cost optimal level of armchair production.

142. At those levels of employments, the least cost amount of armchairs produced will be:

4.5. Returns to Scale

In the long-run, supply of both labor and capital becomes elastic. Firms can therefore employ more of both labor and capital to increase their production. We now turn to determine the amount by which output changes if a firm increases all its inputs proportionately (by the same

percentage). This is what we mean by returns to scale. It shows the percentage change in output as the firm changes all its inputs *by the same proportion*.

We have three types of returns to scale.

CXLIII. Increasing Returns to Scale (IRS):

If output rises more than in proportion to an equal percentage increase in all inputs the production function is said to exhibit increasing returns to scale (IRS).

Example: A given production function exhibits IRS if doubling all inputs more than doubles the output: $f(aL, aK) > af(L, K)$, $a>1$.

Reason for IRS:

Although a firm could duplicate a small factory and double its output, the firm might be able to more than double its output by building a single large plant, allowing for greater specialization of labor or capital. In the two smaller plants, workers have to perform many unrelated tasks such as operating, maintaining, and fixing the machines they use. In the large plant, some workers may specialize in maintaining and fixing machines, thereby increasing efficiency. Similarly, a firm may use specialized equipment in large plant but not in small plant.

CXLIV. Decreasing Returns to Scale (DRS):

If output rises less than in proportion to an equal percentage increase in all inputs, the production function exhibits decreasing returns to scale (DRS).

Example: Doubling all inputs less than doubles output: $f(aL, aK) < af(L, K)$, $a>1$.

Reasons for DRS:

- Difficulty of organizing, coordinating and integrating activities increase with firm size. An owner may be able to manage one plant well but may have trouble of running two plants.

- Large team of workers may not function as well as small teams, in which individuals take greater personal responsibility.

CXLVII. Constant Returns to Scale (CRS):

If output rises by the same proportion to an equal percentage increase in all inputs, the production function exhibits constant returns to scale.

Example: When all inputs double, output also doubles: $f(aL, aK) = af(L, K)$, $a>1$.

Review question

Part I. Multiple Choice Items

148. Which one of the following is **true** about marginal product?
 - a. It is the measure of how outputs affect inputs.
 - b. It lets producers know how much a product is going to cost.
 - c. It describes the relationship between average total cost and marginal cost.
 - d. *It is the measure of impact of an added unit of input on output.*
2. In the second stage of short-run production function,
 - a. There is negative return to a variable input.
 - b. *TP is increasing at a decreasing rate.*
 - c. AP is increasing throughout.

- d. MP is greater than AP.
 - e. MP is increasing throughout.
3. When $APL > MPL$,
- a. APL is falling.
 - b. The firm is producing in the second stage of production.
 - c. The firm is producing in the third stage of production.
 - d. All are correct.
4. The very reason for the existence of **negative marginal returns** in production is:
- a. Mere presence of too much variable inputs.
 - b. Presence of too much fixed inputs.
 - c. *Presence of too many variable inputs as compared to fixed inputs.*
 - d. None of the above
5. The optimum of the firm in production in the long-run is attained at the point where:
- a. An isocost line intersects an isoquant curve.
 - b. Marginal rate of technical rate of substitution is equal to the ratio of input prices.
 - c. An isocost line is tangent to isoquant curve.
 - d. All of the above.
 - e. Only B and C.

Part II. Work Out Items

153. Assume a firm with a production function given by and with a limited outlay of *15,000 birr*. If input prices are as $w=500$ *birr* and $r=1500$ *birr*, answer the questions that follow.

1. Determine following functions: APL , APK , MPL and MPK .
2. Find and .
3. How much L and K must be acquired at the optimum of the producer?

How much is the maximum possible output at that cost of production?

157. A firm has the production function. The slope of the firm's isoquant at the point $(x, y) = (70, 50)$ is _____.
 8. suppose the production function for potato chips is a Cobb-Douglas production function that can be written:

$$QPC = 200K \cdot 6L^4$$

Where QPC = number of bags of potato chips produced per hour, K is the number of deep kettle frying machines employed per hour and L is the number of persons working per hour.

- a. Derive expressions for the marginal product of capital and the marginal product of labor in this particular production process.
- b. You, as plant manager, are currently leasing 100 frying machines and employing 3000 employees. Distributors pay you \$.65 per bag for the chips. At the moment, no more deep kettle frying machines are available for leasing but you can choose to hire more if they will add to your profits. If labor is currently receiving a wage of \$12.00 per hour, would you consider hiring one more person? More than one?
 Explain your reasoning. (To frame the question, remember that an extra person would generate additional revenue via the extra product they produce, and additional cost via the wage they are paid.)
- c. Several months have passed and it is once again possible to release more deep kettle frying machines if you choose to do so. Briefly describe the information you will need to make that decision.

In the previous chapter you have discussed the nature of production. This chapter will discuss about the nature of cost in order to understand the supply side of the market in a more rigorous way.

5.1 Cost Function and Some Types of Costs

Social cost versus private cost

The cost of producing a good can be seen from two perspectives. These perspectives are the society and the individual producer. The social cost of producing a good measures the cost incurred by all society. Social cost takes in to account the cost incurred by those who are directly involved and those who are indirectly affected by the production of the good in question. In other word, social cost takes into account all negative externalities due to the product. Therefore, social cost will be given by the sum of producer's cost and external cost.

This means; **Social cost= Producer's cost + external cost**

Producer's cost includes all expenses the producer incurs in order to produce the good. However, if the cost has to be seen from the society's perspective external agents can incur some cost. This cost can come, for example in the form of pollution, which may damage the interest of the society by causing health problems or by decreasing the productivity of inputs. Therefore, if one wants to measure the full cost of a commodity to the society all negative externalities have to be taken in to account.

Social costs are usually considered in implementing for non – profit projects aiming at improving the welfare of the society. However, some private projects may also consider social costs.

While social costs attempts to see how much the society as a whole pay for a commodity, private costs looks at the costs from the individual's perspective.

In measuring private cost, no attention is given to external costs. Therefore, private cost measures the cost incurred by the individual producer to produce goods and services. So we can say private cost is the producer's cost.

This chapter concerns about the cost incurred by the firm. Therefore, we will limit ourselves to the private cost.

Though, all producers consider the cost incurred by them, ignoring the external cost, there is no uniformity in measuring private cost. This will lead us to the discussion on accounting and economic cost.

Accounting versus economic cost

Accounting cost refers to the money expenses incurred by a firm in producing a commodity. In other words, accounting cost is the monetary value of all purchased inputs used in production.

Accounting cost ignores the cost of non – purchased (self – owned) inputs. It takes in to account the cost of purchased input only. In the real world economy, entrepreneurs may use some resources, which may not have direct monetary expense, since the entrepreneur can own these inputs himself or herself.

Accounting cost takes only those expenses which are direct such as wage and salaries of hired laborers, cost of raw materials, depreciation allowances, interest on borrowed funds, utility expenses such as electricity, water supply, telephone, etc. These costs are said to be explicit cost.

Explicitly costs are out of pocket expenses for the purchased inputs.

Therefore, we can say accounting cost considers only the explicit costs by ignoring the cost of non – purchased inputs.

If a producer calculates her cost by considering only the costs incurred for purchased inputs, then her profit will be accounting profit.

Therefore; Accounting profit = Total Revenue – Accounting cost=Total Revenue – Explicit cost

On the other hand, the economic cost of producing a commodity considers the monetary value of all inputs – purchased and non – purchased. Calculating economic costs will be difficult, since there are no direct monetary expenses for non – purchased inputs.

The monetary value of non – purchased inputs is obtained by estimating the monetary cost by considering the opportunity cost of the non – purchased input. This means, we try to calculate the sacrifice we have made by using the resources in the firm and not in the second best alternative employments.

For example, if Mr. X quit a job which pays him Birr 1500.00 per month in order to run a firm he has established, then the opportunity cost of his labor's will be taken to be Birr 1500.00 per month (the salary he has forgone in order to run his own business).

Another example, if Mrs. Y, withdrew Birr 50,000 from her saving account which she used it to purchase raw materials for her factory. Then, though nobody will ask her to pay interest since the money is her own, this does not mean, the money she has used is a free resource. It has a cost and the cost can be the forgone interest income she has sacrificed. Since, the bank will no more pay her interest on the money she has withdrawn. If for example, the saving interest rate is 4% per year, then Mrs. Y would sacrifice an interest income of Birr 2000 per year. This will be the opportunity cost of using the money in her business.

So, in calculating economic cost by making such estimations, it is possible to get the monetary value of all inputs. The estimated opportunity cost is not direct expenses like explicitly cost but they are indirect or implicit

The estimated monetary cost for non –purchased inputs is known as implicit cost.

Economic cost, therefore, is given by the sum of implicit cost and explicit cost. Firms measuring their cost in economic way will have economic profit which is given us.

Economic profit =Total Revenue – Economic cost= Total Revenue – Explicit cost – Implicit cost

Economic profit will give the real profit of the firm since all costs are taken in to account. Accounting profit of a firm will be greater than economic profit by the amount of implicit cost. If all inputs are purchased from the market accounting and economic profit will be the same. But, if implicit costs exist, then accounting profit will be larger than economic profit.

In the theory of cost we assume costs to be measured in economic cost way in which the cost of all inputs will be taken in to account.

5.2. Short – Run Cost Function

Cost function

A cost function shows the total cost of producing a given level of output. A cost function can be described using equations tables or curves. A cost function can be represented using an equation as follows.

$C = f(Q)$ Where C = total cost of production

Q = the level of output.

Short – run total costs

In the short – run total cost (TC) can be broken down in to two – total fixed cost (TFC) and total variable cost (TVC).

Total Fixed cost (TFC): Total fixed cost refers to that part of the total cost which does not change as output changes. It is the cost of fixed inputs used in production. The most commonly cited examples include rent for land and building, interest on borrower funds, insurance expenses, salaries of permanent staffs, etc.

Total variable cost (TVC): Total variable cost (TVC) is that part of the total cost that varies with the level of output. This means TVC depends on the level of output. TVC is the cost incurred for the variable inputs used in the production. The most commonly cited examples include cost of raw materials, cost of fuel, expenses for utilities such as power consumption, water, telephone, wages of temporary workers, etc.

The sum of TFC and TVC will always be equal to TC in the short –run. In short this can be written as:

$$TC = TFC + TVC$$

Average costs in the short –run

From the total costs, it is possible to derive their respective average costs.

Average total cost (average cost): Average total cost is the total cost per unit of output. It is calculated by dividing the total cost for the level of output symbolically;

$$AC = TC/Q \dots \text{Where } AC = \text{Average cost}$$

TC = Total cost

Q = Level of output

Average Fixed cost (AFC): Average fixed cost is total fixed cost per unit of output. It is calculated by dividing TFC for the level of output.

Symbolically; $AFC = TFC/Q \dots \text{Where } AFC = \text{Average fixed cost}$

TFC = Total fixed cost

Q = Level of output

Average variable cost (AVC): Average variable cost is total variable cost per unit of output. It is calculated by dividing total variable cost for the level of output.

Symbolically; $AVC = TVC/Q \dots \text{Where } AVC = \text{Average variable cost}$

TVC = Total variable cost

The short –run average costs have a relationship among them. We can derive this relationship by starting with the total cost.

$$TC = TFC + TVC$$

Dividing both sides by output (Q) will give

$$TC/Q = TFC/Q + TVC/Q \text{ or:}$$

$$AC = AFC + AVC$$

Therefore, AC is given by the sum of AFC and AVC

Marginal cost (MC)

Marginal cost refers to the additional cost incurred in order to produce one additional output. It can be also be defined as change in total cost per unit change in output. Symbolically,

$\dots \text{Where } MC = \text{marginal cost}$

TC = Change in total cost

Q = Change in output

From a total cost function, $C=f(Q)$, the marginal cost will be given by the first derivative of total cost function with respect to output. Symbolic,

$$\text{Given } C = f(Q)$$

$$MC = dC/dQ$$

It is also possible to get MC from total variable cost (TVC). This can be seen as follows. We defined MC as:

, But $TC = TFC + TVC$, so

but since TFC is constant. This implies that;

Therefore, .

This implies MC is the change in either TC or TVC per unit change in output. This means TFC does not affect the marginal cost.

To examine the behavior of the short – run costs and understand their relationship, let us take the following typical short – run cost function.

Output	TFC	TVC	TC	AFC	AVC	AC	MC
1	2	3	4	5	6	7	8
0	60	0	60	-	-	-	-
1	60	15	75	60	15	75	15
2	60	25	85	30	12.5	42.5	10
3	60	30	90	20	10	30	5
4	60	50	110	15	12.5	27.5	20
5	60	80	140	12	16	28	30
6	60	125	185	10	25.3	35.3	45

Table 5.1 Short-run cost Schedule

As you can see from the table, the TFC is 60, which is a constant and remains constant at all levels of output. However, TVC and TC continuously rise as output increase

The difference between TC and TVC at all levels of output is constant 60 which is the fixed cost. At output level zero, TVC is zero but TC has the value of the TFC .

AFC continuously declines as output increases. AVC , AC and MC have the same change of variation all first decrease; reach their respective minimum and then increase.

We can sketch the short – run cost curves as follows.

TC,

TVC

Output (Q)

ATC MC AC

AFC AVC

AVC

MC

AFC

Fig 5.1 Short – run cost curves.

Out put

TFC is represented by a horizontal line, since its value is constant at all levels of output.

TC and TVC are represented by continuously rising curves they are parallel since they have the same slope given by MC. They both increase continuously first by a decreasing then by an increasing rate which is the result of the behaviors of the MC. However, there is a constant gap between TC curve and TVC curve and this gap is the fixed cost.

- When $Q = 0$

$$TVC = 0 \text{ and } TC = TFC$$

AFC declines continuously and approaches both axes asymptotically. AFC is rectangular hyperbola. AVC, AC and MC curves have 'U' shape. They first decline reaches their respective minimum and then rises. MC curve cuts the AVC curve and the AC curve at their respective minimum point. This is because as long as MC is below AC, it pulls the averages down. When MC is above AC or AVC, it pulls the averages up. When MC equals AC it is neither falling nor rising (i.e. AC or AVC are at their minimums). The gap between AC curve and AVC curve diminishes as output increases because the

difference between them is AFC and it declines as output increase. So long as AFC and AVC fall, AC will also fall because $AC = AFC + AVC$

Example

Suppose the short – run cost function of a firm is given by: $C=2Q^3 - 2Q^2 + Q + 10$. Find

- d. The expressions for TFC & TVC
- e. The expressions for AFC, AVC & AC and MC
- f. The minimum values of MC and AVC.

Solution

Given $C=2Q^3 - 2Q^2 + Q + 10$

159. $TFC = TC$ at $Q=0$ Therefore ,

$$TFC = 10$$

$$TVC = TC - TFC = 2Q^3 - 2Q^2 + Q + 10 - 10$$

$$TVC = 2Q^3 - 2Q^2 + Q$$

$$160. \quad AFC = TFC/Q = 10/Q$$

$$AVC = TVC/Q = (2Q^3 - 2Q^2 + Q)/Q = 2Q^2 - 2Q + 1$$

$$AC = TC/Q = (2Q^3 - 2Q^2 + Q + 10)/Q = 2Q^2 - 2Q + 1 + 10/Q$$

$$MC = dC/dQ = 6Q^2 - 4Q + 1$$

161. Minimum value of MC is obtained at output level, which will make the first derivative of MC zero. Therefore, to find the minimum value of MC:

$$dMC/dQ = 12Q - 4 = 0$$

$$Q = 1/3$$

To find the minimum value of $MC = 6(1/3)^2 - 4(1/3) + 1 = 0.37$

We follow the same step to get the minimum value of AVC

$$AVC = 2Q^2 - 2Q + 1$$

$dAVC/dQ = 4Q - 2 = 0$, $Q = 0.5$. To find the minimum value of AVC

$$AVC = 2(0.5)^2 - 2(0.5) + 1$$

$$= 0.5 - 1 + 1$$

= 0.5

5.2.1. The relation between production and cost.

As we know production involves cost. Therefore, we have to analyze how production and cost are related

In this section we will see how marginal product of labor and marginal cost are related, and also how average product of labor and average variable cost are related.

Suppose a firm in the short – run uses labor as a variable input and capital as a fixed input. We can put the production function as follows.

$$Q = f(L, K) \dots \text{Where } Q = \text{level of output}$$

L = variable input (labor)

K = fixed input (capital)

Let the price of the variable input labor be given by w , which is constant. Using the production function it is possible to show the relations between production cost variables.

a) How MC and MPL are related?

Let us start by defining the MC of this firm as:

But $TVC = w \cdot L$

, but Therefore,

From the above relation we can clearly see the relationship between MC and MPL.

The above expression shows that MC and MPL are inversely related, which means when initially MPL increases, MC decreases and when MPL is at maximum MC must be at a minimum and when finally MPL declines MC increases.

B. How average variable cost is related with average product of labor?

, but $TVC = w \cdot L$, so

but

Therefore,

This relation also shows inverse relation between AVC and APL which means when APL increases AVC decreases & when APL is at a maximum AVC is at a minimum and when finally APL declines AVC increases.

From the above relations we can conclude that the reason for the 'U' shape of the MC of AVC is the law of variable proportion which causes the MPL to decline.

We can also see the relationship between these production and cost variables using graphs. This is given as follows.

MPL

Fig 5.2 Production and cost curves

From the above diagram we can conclude that the MC curve is the mirror image of MPL curve and AVC curve is the mirror image of APL curve

5.3. Long – Run Cost Analysis.

The long – run refers to the fact that economic agents – consumers and managers – can plan ahead and choose many aspects of the “short – run” in which they will operate in the future. Thus in a sense, the long – run consists of all possible short – run situations among which an economic agent may choose.

In short, we can say an economic agent operates in the short – run and plans in the long run. To see how the long – run average cost curve (LACC) is derived from the short – run average cost curves (SACC), suppose technology is such that plants in a certain industry can have only three different sizes. That is, the fixed capital equipment comprising the “plant” is available in only three sizes – small, medium and large.

Let the three plants be represented by their average cost curves SAC1, SAC2 and SAC3 as shown in Figure 5.3 below.

Cost SAC3

Figure 5.3 Short –run Average Cost Curves

Each curve represents the scale of the firm SAC depicts small scale while SAC2 and SAC3 depicts medium and large scale. In the long – run, a manager has to choose among the three scale of operation available as represented by the three short – run average cost curves. The manager's choice of plant will depend on planned or expected output. Thus, if the firm plans to produce output OQ1 the small plant will be selected. If expected output is OQ2, the medium plant is selected, and so forth. Such decisions would be made because the manager chooses the plant capable of producing the planned output at the lowest unit cost.

If planned output is either OQ11, or OQ21, the decision in evolves some ambiguity. At each of these points, two plants have the same average cost.

A manager might select the larger plant in order to meet a possible expansion of demand. On the other hand, the smaller plant would be preferred if a contraction of demand were possible. In these two cases, the decisions could involve consideration others than least – cost output.

Suppose planned output is OQ1. There is no ambiguity in this case and the plant represented by SAC1, will be built. Now suppose that it actually turns out to be desirable to produce OQ2 units. This can be done with the existing plant at an average cost of OC₁ per unit. In the short – run, there are no other options and this is the only way to produce OQ2 units. But the entrepreneurs can plan for the future. When the old plant has "worn out", it can be replaced with a new one – and it will be a medium – size plant because OQ2 can be produced for an average cost of OC₂ per unit, substantially less than with the small plant.

If the firm expands its scale by the three plants represented by SAC1, SAC2 and SAC3 curves, the thick wave like portion of these curves form the long – run average cost (LAC) curve.

The dotted portions of these SAC curves are of no consideration during the long – run because the firm would change the scale of plant rather than operate on them.

However, the long – run average cost curve fitted to the SAC curves so that it is tangent to each of them at some point, as shown in Fig 5.4 below, where SAC1, SAC2 , SAC3, SAC4 ,SAC5, SAC6 are the short – run curves . The LAC is also known by the name of the "envelope curve "since it envelopes

the short – run average cost curves.

Cost per unit

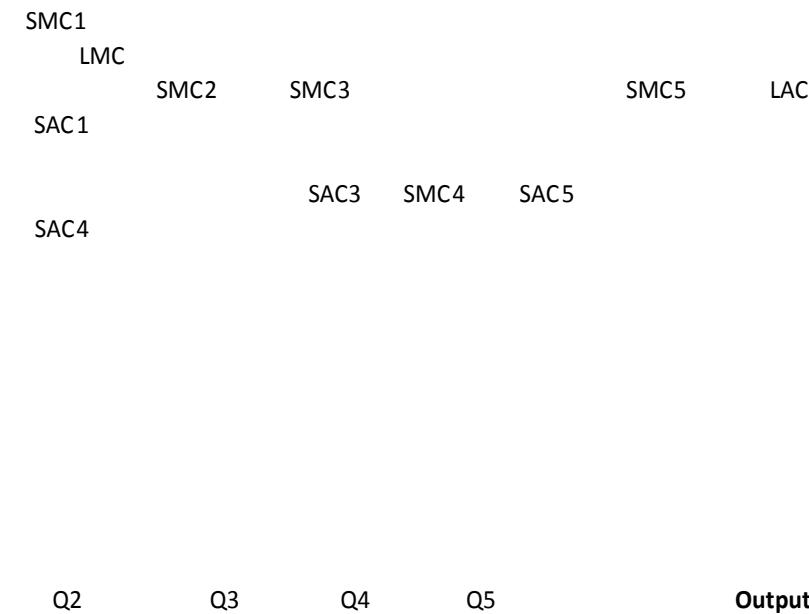


Fig 5.4 Average and Marginal cost curves

These many curves generate LAC curve as a planning device. Suppose the entrepreneurs think the output associated with point A in Fig 5.4 will be most profitable. The plant represented by SAC1 will be built, because it will allow production of this output at the least possible cost per unit.

For average cost given by SAC1 unit cost would be reduced by expanding output to the amount associated with point B (Q2), the minimum point on SAC1. If demand conditions were suddenly changed so that this larger output was desirable, the entrepreneur easily expands and would add to profitability by reducing unit cost. Nevertheless, when setting future plans, the entrepreneurs would decide to construct the plant represent by SAC2 because the firm could reduce unit costs even more. It would operate at point there by lowering unit cost from the level at point B on SAC1

The long – run planning curve, LAC, is a locus of points representing the least unit cost of producing the corresponding output. The entrepreneurs determined the short – run plant that yields the least unit cost of producing the volume of output desired.

As can be observed from Fig 5.4 the LAC is tangent to the lowest point of curve SAC4 at point M and Q4 is the optimal output at the minimum cost with this optimum plant we have the optimum firm. If the firm produces less than the optimum output Q4 it is not reaping fully economies of production and it produces beyond Q4 it faces diseconomies of scale. In both cases, the plants such as SAC1, SAC2, SAC3, SAC5 & SAC6 have higher average cost of production than plant SAC4. In other words, the downward trend in the LAC indicates that until the output reaches the level of Q4, the firm is of non-optimal size. Similarly, expansion of the firm beyond production capacity Q4 results into increase in SAC as well as LAC. ***It follows that, given the technology, a firm trying to minimize its average cost over time must choose a plant size which gives the minimum LAC where SAC = SMC = LMC. This size of the plant assures most efficient utilization of the resources. Any change in output level (increase or decrease) will make the firm enter the area of non – optimality.***

There is also long –run marginal cost curve shown in Fig 4.4 the line labeled as LMC. It indicated the extra costs entailed in enlarging production facilities so as to get greater quantities of output. It bears precisely the same relationship to the LAC curve that the SMC curve bears to its SAC curve.

Review Questions

PART I: Multiple choice questions

----- 1. Which one of the following statement is true?

- a. Accounting cost can never be greater than economic cost
- b. Explicit costs are the estimated cost of non – purchased inputs
- c. Accounting costs ignore the cost of purchased inputs
- d. Accounting profit can be smaller than economic profit

----- 2. Implicit costs are costs_____.

- 162. having direct expenses
- 2. which are not incurred by the firm
- 3. of non – purchased inputs
- 4. which are not part of private cost

----- 3. In the short – run period of production, when output increases, the AFC

- 166. Increase
- 2. Decreases
- 3. Remain the same
- 4. None of the above

----- 4. MC is given by all of the following except_____

- a. The slope of total product curve
- b. The slope of TVC curve
- c. The slope of TC curve
- d. None of the above

----- 5. All of the following curves are U – shaped except

- 170. LAC b. AFC c. AVC d. MC

----- 6. Suppose a firm pays Br.60 to each unit of the variable input used and if the marginal product of labor is 3 units, then the marginal cost will be -----

- 171. Birr.20 b. Birr.120 c. Birr.40 d. None of the above

----- 7. In the short – run, if the firm does not produce any output ($Q = 0$), then-----.

- a. Total cost is zero
- b. Variable cost is zero
- c. Total cost and total fixed cost will be equal

d. All except A

----- 8. Which one of the following cost is more likely to be a variable cost of the firm?

- a. Insurance payment
- b. Rent for land
- c. Costs of raw materials
- d. Interest on borrowed funds.

----- 9. The reason for the eventual rise of LAC is -----.

- e. Economies of scale
- f. Diseconomies of scale
- g. Over – specialization
- h. All of the above

----- 10. Which one of the following is true at the minimum point of LAC?

- i. $LAC = LMC > SAC = SMC$
- j. $LAC = SAC < SMC = LMC$
- k. $LAC = LMC = SAC = SMC$
- l. $LAC = LMC = SAC > SMC$

PART II: Discussion questions

- m. Compare and contrast the following cost concepts.
 - i. Social cost and private cost
 - ii. Explicit cost and implicit cost
 - iii. Economic and accounting cost
- n. Distinguish between short – run and long – run cost functions
- o. Why does long – run average cost curves fall, and then rise? Why do short – run average cost curves first fall, and then rise?
- p. What happens to average total cost as:
 - i. Marginal product increases?
 - ii. Average product increases?
- q. When does average cost increase?

Answer this in terms of:

- i. The relation of average cost to marginal cost
- ii. The relation between the increase in AVC and the decrease in AFC.

PART III: Workout Questions

1. Fill in the blanks in the following table.

Output (Q)	TC	TFC	TVC	AFC	AVC	AC	MC
0							
1					30		
2			40				
3		10			15		
4	65						
5							15

2. Given a short run cost function as.

, Find the minimum value of AVC and MC

CHAPTER SIX: MARKET STRUCTURES

In the previous chapters, we have noticed that the overriding objective of the firm is profit maximization. In this chapter we are going to discuss the profit maximizing conditions of a firm under various product markets. Economists categorize industries into four distinct market structures: pure competition, pure monopoly, monopolistic competition and oligopoly. These four market models differ in several respects: the number of firms in the industry, whether those firms produce a standardized product or try to differentiate their products from those of other firms, and how easy or how difficult it is for firms to enter the industry. In terms of competitiveness, the spectrum of market structures reaches from pure competition, to monopolistic competition, to oligopoly, to pure monopoly

6.1. Pure Competition

6.1.1. Definition

Pure competition involves a very large number of firms producing a standardized product (that is, a product identical to that of other producers such as corn or cucumbers). New firms can enter or exit the industry very easily.

Although pure competition is relatively rare in the real world, this market model is highly relevant. A few industries more closely approximate pure competition than any other market structure. In particular, we can learn much about markets for agricultural goods, fish products, foreign exchange, basic metals, and stock shares by studying the pure-competition model. Also, pure competition is a meaningful starting point for any discussion of price and output determination. Moreover, the operation of a purely competitive economy provides a standard, or norm, for evaluating the efficiency of the real-world economy.

6.1.2. Characteristics/Assumption of the Model

The model of perfect competition was constructed based on the following assumptions or imaginations.

3. Large number of sellers and buyers. A basic feature of a purely competitive market is the presence of a large number of independently acting sellers, often offering their products in large national or international markets.

Perhaps the closest we come to perfect competition is in the market for such agricultural commodities as wheat and cotton. There, we may have a large number of producers each too small to affect commodity price.

4. Standardized products. Purely competitive firms produce a standardized (identical or homogeneous) product. This means a commodity produced by one firm is the same as that produced by another firm and. As long as the price is the same, consumers will be indifferent about which seller to buy the product from. Buyers view the products of firms B, C, D, and E as perfect substitutes for the product of firm A. Because purely competitive firms sell standardized products, they make no attempt to differentiate their products and do not engage in other forms of non-price competition (i.e. advertisement).

5. A firm is price-taker. In a purely competitive market, individual firms exert no significant control over product price. Each firm produces such a small fraction of total output that increasing or decreasing its output will not perceptibly influence total supply or, therefore, product price. In short, the competitive firm is a **price taker** and can sell any amount of the commodity at the prevailing market price. It cannot change market price; it can only adjust to it. That means that the individual competitive producer is at the mercy of the market.

6. Free entry and exist. New firms can freely enter and existing firms can freely leave purely competitive industries. No significant legal, technological, financial, or other obstacles prohibit new firms from selling their output in any competitive market.

7. Perfect Knowledge of Market Conditions. The buyers and sellers have complete information about market conditions. The buyers have perfect knowledge about the conditions prevailing in the market for purchasing a commodity at the market price. If they do not have perfect knowledge, they may pay a higher price for the product, and then there would be different prices for the same product. So the buyers and sellers must have perfect knowledge to be able to buy goods and raw materials at the prevailing market price.

8. Absence of government regulation/intervention. By assumption, there is no government intervention in the market. That is there is no tax, subsidy etc. price and output is determined in the market through the interaction of demand and supply.

6.1.3. Demand and Revenue Functions under Perfect Competition

Because each purely competitive firm offers only a negligible fraction of total market supply, it must accept the price predetermined by the market; it is a price taker, not a price maker. The demand schedule/function faced by the *individual firm* in a purely competitive industry is perfectly elastic at the market price.

Graphically,

Fig 6.1. The demand curve that a perfectly competitive firm faces is horizontal line with intercept at the market price. This indicates that sellers sell any quantity demanded at the ongoing market price and buyers buy any amount they want at the ongoing market price.

Total, average and marginal revenues

Total Revenue (TR) - Given the horizontal demand function at the ongoing market price, the total revenue of a firm operating under perfect competition is given by the product of the market price and the quantity of sales, i.e.,

$$TR = P \cdot Q$$

Since the market price is constant at P^* , the total revenue function is linear and the amount of total revenue depends on the quantity of sales. To increase his total revenue, the firm should sell large quantity.

Graphically, the TR curve is as shown below.

Fig 6.2 The total revenue of firm operating in a perfectly competitive market is linear (and increasing function) of the quantity of sales.

Average Revenue (AR) - It is the TR divided by the quantity of sales. i.e. Hence, $AR = P$. The average-revenue (AR) curve also coincides with the demand curve.

Marginal Revenue (MR) - MR is the change in total revenue that occurs when one more unit of the output is sold, i.e..Hence $MR=P$

Note: The marginal revenue (MR) and average revenue (AR) of a firm operating under perfect competition are equal to the market price (demand curve). $P=MR=AR$.

6.1.4. Short Run Equilibrium of the Firm

Because the purely competitive firm is a price taker, it can maximize its economic profit (or minimize its loss) only by adjusting its *output*. And, in the short run, the firm has a fixed plant. Thus it can adjust its output only through changes in the amount of variable resources (materials, labor) it uses. It adjusts its variable resources to achieve the output level that maximizes its profit. Under perfect competition, the firm is said to be in equilibrium when it produces that level of output which maximizes its profit, given the market price.

There are two ways to determine the level of output at which a competitive firm will realize maximum profit or minimum loss. One method is to compare total revenue and total cost; the other is to compare marginal revenue and marginal cost. Both approaches apply to all firms, whether they are pure competitors, pure monopolists, monopolistic competitors, or oligopolists.

Total Approach (TR-TC approach)

In this approach, a firm maximizes total profits in the short run when the (positive) difference between total revenue (TR) and total costs (TC) is greatest.

TC TR

Graphically

TC,TR

Q

Q0 Q_e Q1

Fig: 6.3. The profit maximizing output level is Q_e because it is at this output level that the vertical distance between the TR and TC curves (or profit) is maximum.

For all output levels below Q_0 and above Q_1 profit is negative because TC is above TR.

Marginal Approach (MR-MC)

In general, it is more useful to analyze the short-run behavior of the firm by using the marginal-revenue–marginal-cost approach. In this approach, the firm compares the amounts that each *additional* unit of output would add to total revenue and to total cost. In other words, the firm compares the *marginal revenue* (MR) and the *marginal cost* (MC) of each successive unit of output. In the short run, the firm will maximize profit or minimize loss by producing the output at which marginal revenue equals marginal cost (as long as producing is preferable to shutting down). More specifically, the perfectly competitive firm maximizes its short-run total profits at the output when the following two conditions are met:

- $MR = MC$
- The slope of MC is greater than slope of MR; or MC is rising).

Graphically, the marginal approach can be shown as follows.

MC, MR

MC

MR

Q* Qe

Fig 6.4: The profit maximizing output is Q_e , where $MC=MR$ and MC curve is increasing. At Q^* , $MC=MR$, but since MC is falling at this output level, it is not equilibrium output. For all output levels ranging from Q^* to Q_e the marginal cost of producing additional unit of output is less than the MR obtained from selling this output. Hence the firm should produce additional output until it reaches Q_e .

Whether the firm in the short-run gets positive or zero or negative profit depends on the level of ATC at equilibrium. Thus depending on the relationship between price and ATC, the firm in the short-run may earn positive profit or normal profit or incur loss and decide to shut-down business.

T. Positive profit - If the ATC is below the market price at equilibrium, the firm earns a positive profit equal to the area between the ATC curve and the price line up to the profit maximizing output (see fig. 6.5 below).

Fig. 6.5. The firm earns a positive profit because price exceeds AC of production at equilibrium

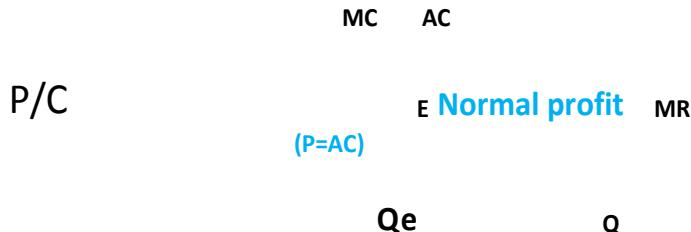
U. Loss - If the ATC is above the market price at equilibrium, the firm earns a negative profit (incurs a loss) equal to the area between the ATC curve and the price line.(see fig 6.6 below).

AC

P

Fig 6.6 A firm incurs a loss because price is less than AC of production at equilibrium.

V. Normal Profit (zero profit) or breaking even point - If the ATC is equal to the market price at equilibrium, the firm gets zero profit or normal profit.

**Fig 6.7** A firm gets a normal profit at a point where price is equal to AC of production at equilibrium.

W. Shutdown point - The firm will not stop production simply because AC exceeds price it incurs a loss in the short-run. The firm will continue to produce irrespective of the existing loss as far as the price is sufficient to cover the average variable costs. This means, if P is larger than AVC but smaller than AC, the firm minimizes total losses. But if P is smaller than AVC, the firm minimizes total losses by shutting down. Thus, $P = AVC$ is the shutdown point for the firm.

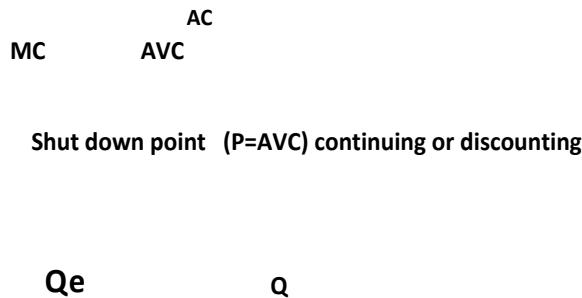


Fig 6.8. A shut down point where the price is equal to the average variable cost of a firm.

Example 1:

Suppose a firm has a TFC of \$2,000, a TVC of \$ 5,000 and a TR of \$6,000 at equilibrium. Should the firm stop its operation? Why?

In fact the firm is incurring a loss of \$ 1,000 because TC ($2,000 + 5,000 = 7,000$) is greater than the total revenue. But the firm should continue production because the TR is greater than TVC. If the firm stops operation, it will lose the fixed cost (\$ 2,000). But if it continues production the loss is only \$ 1,000 (TR-TC). Thus, the firm requires a minimum TR of \$ 5,000 to continue operation. If the TR is equal to \$ 5,000, the firm is indifferent in between choosing to continue or to discontinue its operations because in both cases the loss is equal to fixed costs. Thus the level output at which TR and TVCs are equal is called shut down output level. In other words, shut down point is the point at which AVC equals the market price.

Example 2

Now let us see how to determine the short run equilibrium of a firm operating in a perfectly competitive market by using a hypothetical example.

Suppose that the firm operates in a perfectly competitive market. The market price of his product is \$10. The firm estimates its cost of production with the following cost function:

$$TC = 10q - 4q^2 + q^3$$

- i. What level of output should the firm produce to maximize its profit?
- ii. Determine the level of profit at equilibrium.
- iii. What minimum price is required by the firm to stay in the market?

Solution

Given: $p = \$10$

$$TC = 10q - 4q^2 + q^3$$

- iv. The profit maximizing output is that level of output which satisfies the following condition

$$MC = MR \text{ &}$$

MC is rising

Thus, we have to find MC & MR first

(MR in a perfectly competitive market is equal to the market price. Hence, $MR = 10$

Alternatively, where $TR = P \cdot q = 10q$

Thus,

$$MC =$$

(To determine equilibrium output just equate MC & MR

And then solve for q .

$$10 - 8q + 3q^2 = 10$$

$$-8q + 3q^2 = 0$$

$$q(-8 + 3q) = 0$$

$$q = 0 \text{ or } q = 8/3$$

Now we have obtained two different output levels which satisfy the first order (necessary) condition of profit maximization

i.e. 0 & 8/3

(To determine which level of output maximizes profit we have to use the second order test at the two output levels i.e. we have to see which output level satisfies the second order condition of increasing MC.

(To see this first we determine the slope of MC

$$\text{Slope of MC} = -8 + 6$$

(At $q = 0$, slope of MC is $-8 + 6 (0) = -8$ which implies that marginal cost is decreasing at $q = 0$. Thus, $q = 0$ is not equilibrium output because it doesn't satisfy the second order condition.

(At $q = 8/3$, slope of MC is $-8 + 6 (8/3) = 8$, which is positive, implying that MC is increasing at $q = 8/3$

Thus, the equilibrium output level is $q = 8/3$

B) Above, we have said that the firm maximizes its profit by reducing $8/3$ units. To determine the firm's equilibrium profit we have calculate the total revenue that the firm obtains at this level of output and the total cost of producing the equilibrium level of output.

$$TR = \text{Price} * \text{Equilibrium output}$$

$$= \$10 * 8/3 = \$80/3$$

TC at $q = 8/3$ can be obtained by substituting $8/3$ for q in the TC function, i.e.,

$$TC = 10(8/3) - 4(8/3)^2 + (8/3)^3 = 23.12$$

Thus the equilibrium (maximum) profit is

$$\pi = TR - TC$$

$$= 26.67 - 23.12 = \$3.55$$

c) To stay in operation the firm needs the price which equals at least the minimum AVC. Thus to determine the minimum price required to stay in business, we have to determine the minimum AVC.

AVC is minimal when derivative of AVC is equal to zero

That is: $= 0$

Given the TC function: $TC = 10q - 4q^2 + q^3$, there is no fixed cost i.e. TC is equal to the TVC.

$$\text{Hence, TVC} = 10q - 4q^2 + q^3$$

$$\text{AVC} = 10 - 4q^2 + q^2$$

(

$$= -4 + 2q = 0$$

($q = 2$ i.e. AVC is minimum when output is equal to 2 units.

The minimum AVC is obtained by substituting 2 for q in the AVC function i.e., Min AVC = $10 - 4(2) + 22 = 6$

Thus, to stay in the market the firm should get a minimum price of \$ 6.

6.1.5. The Short Run Supply Curve of the Firm and the Industry

Since the perfectly competitive firm always produces where $P = MR = MC$ (as long as P exceeds AVC), the firm's short-run supply curve is given by the rising portion of its MC curve over and above its AVC, or shutdown point. The industry-supply curve is the horizontal summation of the supply curves of the individual firms. That is, the total quantity supplied in the market at each price is the sum of the quantities supplied by all firms at that price.

6.1.6. Long- Run Equilibrium of the Firm

In the long run firms already in an industry have sufficient time either to expand or to contract their plant capacities. More important, the number of firms in the industry may either increase or decrease as new firms enter or existing firms leave. That is, the firm is in the long run equilibrium when the market price is equal to the minimum long run AC. Thus since price is equal to long run AC (LAC now on) at the long run equilibrium, firms will be earning just normal profits (zero profits), which are included in the LAC. Firms get only normal profit in the long run due to two reasons.

First, if the firms existing in the market are making excess profits (the market price is greater than their LACs) new firms will be attracted to the industry seeking for this excess profit. The entry of new firms results in two consequences:

A. The entry of new firms will lead to a fall in market price of the commodity (which is shown by the down ward shift of the individual demand curve). This happens because entry of new firms will increase the market supply of the commodity (which is shown by the right ward shift of the industry supply), resulting in the lower market price. Moreover, if firms are getting excess profit, they have an incentive to expand their capacity of production, which increases the market supply and then reduces the market price.

B. Moreover, the entry of new firms results in an upward shift of the cost curves. This happens because, when new firms enter into the market the demand for factors of production increases which exerts an upward pressure on the prices of factors of production. An increase in the price of factors of production in turn shifts the cost curves upward. These changes (decrease in the market price and upward shift of the cost curves) will continue until the LAC becomes tangent to the demand curve defined by the market price. At this time, entry of new firms will stop since there is no positive profit (since $P = LAC$) which attracts new firms in to the market.

Second, if the firms are incurring losses in the long run ($P < LAC$) they will leave the industry (shut down). This will result in higher market price (because market supply of the commodity decreases) and lower costs (because the market demand for inputs decreases as the number of firms in the market decreases). These changes will continue until the remaining firms in the industry cover their total costs inclusive of the normal rate of profit.

Thus, due to the above two reasons, firms can make only a normal profit in the long run.

The following figure shows how firms adjust to their long run equilibrium position excess profit (higher price than minimum lack) if the market price is p , the firm is making excess profit working with plant size whose cost is denoted by SAC, (short run average cost1). It will therefore have an incentive to build new capacity or larger plant size and it moves along its LAC. At the same time new firms will be entering the industry attracted by the excess profits. As quantity supplied in the market increases (by the increased production of expanding old firms and by the newly established ones) the supply curve in the market will shift to the right and price will fall until it reaches the level of P_1 , at which the firms and the industry are in the long- run equilibrium.

Fig 6.9: Long run equilibrium of the firm. Entry of new firms reduces the market price from p to p_1 (in panel A) and the long run equilibrium is established at E (panel B).

The condition for the long run equilibrium of the firm is that the long run marginal cost (LMC) should be equal to the price and to the LAC i.e. $SMC = LMC = SAC = LAC = P = MR$

This implies that at the minimum point of the LAC the corresponding short run plant is worked at its optimal capacity so that the minimum of the LAC and SAC coincide.

6.1.7. Perfect Competition and Optimal Resource Allocation

In the perfect competition, the market mechanism leads to an optimal allocation of resources. The optimality is shown by the following conditions all of which prevail in the long run equilibrium of the industry;

- a) The output is produced at the minimum feasible cost. That is all firms produce at the minimum of their LAC.
- b) Consumers pay the minimum possible price which just covers the marginal cost of production, that is, price equals just opportunity cost so that the consumers are not exploited.
- c) Plants are used at full capacity in the long-run so that there is no waste of resources. That is, at the long run equilibrium the short run average cost is also at minimum.
- d) Firms earn only normal profits.

These conditions justify the fact that perfect competition results in optimal resource allocation.

Review questions

1. MR Curve for a competitive firm is horizontal because:
 - a. product price falls as output increases.
 - b. the law of diminishing marginal utility is at work.
 - c. the market demand for this product is perfectly elastic.
 - d. the firm is a price taker.
2. We know the firm is a price taker because:
 - a. its MC curve slopes upward.
 - b. its ATC curve is U-shaped.
 - c. its MR curve is horizontal.
 - d. MC and ATC are equal at the profit-maximizing output.
3. At the firm's profit-maximizing output:
 - a. total revenue equals total cost.
 - b. it is earning an economic profit.
 - c. allocative, but not necessarily productive, efficiency is achieved.
 - d. productive, but not necessarily allocative, efficiency is achieved.
4. The equality of P , MC , and minimum ATC:
 - a. occurs only in constant-cost industries.
 - b. encourages entry of new firms.
 - c. means that the "right goods" are being produced in the "right ways."
 - d. results in a zero accounting profit.
5. When $P = MC = \text{lowest ATC}$ for individual firms, in the market:
 - a. consumer surplus necessarily exceeds producer surplus.
 - b. consumer surplus plus producer surplus is at a maximum.
 - c. producer surplus necessarily exceeds consumer surplus.

d. supply and demand are identical.

6. Why is the equality of marginal revenue and marginal cost essential for profit maximization in all market structures? Also explain why price can be substituted for marginal revenue in the $MR = MC$ rule when an industry is purely competitive.

7. In the long run, all firms under perfectly competitive conditions earn only normal profit. Why is this true?

8. Suppose that you are given the following information about a particular perfectly competitive industry.

Market Demand: $Q_D = 6500 - 100P$

Market Supply: $Q_S = 1200P$

Firm total cost function: $C(q) = 722 +$

Assume that all firms are identical and that the market is characterized by perfect competition.

- x. Find the market equilibrium price and quantity supplied by each firm.
- y. Determine the profit of a firm
- z. Would you expect entry into or exit from the industry in the long-run? Explain
- a. What is the lowest price at which each firm would sell in the long run?

9. Suppose that a competitive firm's marginal cost of producing output q is given by $MC(q) = 3 + 2q$. Assume that the market price for the firm's product is \$9.

a. What level of output will the firm produce?

b. Suppose the firm's fixed cost is known to be \$3. Will the firm be earning positive, negative or zero profit?

6.2 Pure Monopoly

6.2.1 Definition:

Pure monopoly is a market structure in which one firm is the sole seller of a product or service (for example, a local electric utility). Since the entry of additional firms is blocked, one firm constitutes the entire industry. Thus, it is at the opposite extreme from perfect competition.

6.2.2. Characteristics/Assumptions

Here are the main characteristics of pure monopoly:

1. Single seller and many buyers. A pure, or absolute, monopoly is an industry in which a single firm is the sole producer of a specific good or the sole supplier of a service; the firm and the industry are synonymous.

2. No close substitutes A pure monopoly's product is unique in that there are no close substitutes. The consumer who chooses not to buy the monopolized product must do without it.

3. Price maker The pure monopolist controls the total quantity supplied and thus has considerable control over price; it is a *price maker* (unlike a pure competitor, which has no such control and therefore is a *price taker*). The pure monopolist confronts the usual downward-sloping product demand curve. It can change its product price by changing the quantity of the product it supplies. The monopolist will use this power whenever it is advantageous to do so.

4. Blocked/barrier to entry A pure monopolist has no immediate competitors because certain barriers keep potential competitors from entering the industry. Those barriers may be economic, technological, legal, or of some other type. But entry is totally blocked in pure monopoly.

6.2.3 Sources/Causes of Monopoly

There are many factors that create monopoly and help the monopolists to maintain monopoly power. These include:

b. Ownership of strategic or key inputs. A firm may own or control the entire supply of a raw material required for the production of a commodity. Such firms are not willing to sell the raw materials to another firm. For example, until the second world war, the aluminum Company of America (Aloca) controlled practically the entire supply of Bauxite(the basic raw material necessary for the production of aluminum), giving it almost a complete monopoly in the production of aluminum in the united states. To come to our country, Ambo Mineral Water can be taken as an example. Ambo mineral water has monopolized the natural mineral water.

c. Exclusive knowledge of production technique. Most of the beverage (soft drink) companies such as Coca Cola Company have maintained monopoly power over supply of their product partly due to exclusive knowledge of the ingredient chemicals required for the production of their product.

d. Patents and copyright. Patents and copyrights are government supported barriers to entry. Patents are granted by the government for 17 years as an incentive to investors. Authors of books, artistic works (such as cassette, video, etc) are the best examples of such monopoly

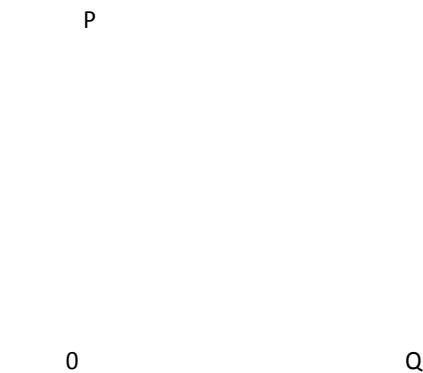
e. Government Franchise and License. Another cause for the emergence of monopoly is government franchise. Franchise is a promise by the government for a firm to prohibit the establishment of another firm (by another person) that produces the same product or offers the same service as the original one.

f. Increasing returns to scale.

6.2.4 Demand and Revenue for the Firm

Under pure monopoly, the firm *is* the industry and faces the negatively sloped industry demand curve for the commodity. As a result, if the monopolist wants to sell more of the commodity, it must lower its price. Thus, the monopolist's MR curve lies below the demand curve, indicating that marginal revenue is less than price (average revenue) at every output quantity but the very first unit.

Graphically, monopoly MR and demand curve is shown as follow



Mathematically, it can be shown that MR is less (steeper) than the AR or demand curve. Suppose a monopolist's demand curve is given by

$$P = a - bQ$$

Where a & b - are any positive constants

P&Q – are price and quantity.

$$TR = P \cdot Q = (a - bQ) Q$$

$$= aQ - bQ^2$$

By definition MR is change in TR that happens due to a one unit change in quantity of sales. Symbolically,

Thus, $MR = (a - 2bQ)$ has a slope which equals twice the slope of demand (average revenue) curves. This implies that MR is less than AR or demand or price.

6.2.5. Profit Maximization

A monopolist seeking to maximize total profit will employ the same rationale as a profit-seeking firm in a competitive industry. If producing is preferable to shutting down, it will produce up to the output at which marginal revenue equals marginal cost ($MR=MC$, and MC is increasing). Depending on the level of AC at this output, the monopolist can have profits, break even, or minimize the short-run total losses. Thus equilibrium demands two conditions to be fulfilled: (i) $MR=MC$ and (ii) slope of MC greater than slope of MR .

Numerical example

Suppose the monopolist faces a market demand function given by $P=40-Q$. The firm has a fixed cost of \$ 50 and its variable cost is given as $TVC=Q^2$ determine:

- g. the profit maximizing unit of output and price
- h. the maximum profit

Solution

Given: $p=40-Q$ $TFC=50$ $TVC=Q^2$

172. equilibrium condition is $MR=MC$, and slope of $MC >$ slope of MR .

$$TR=P.Q = (40-Q) Q = 40Q - Q^2$$

$$TC=TFC+TVC = 50 + Q^2$$

Now,

$$MR=MC \quad 40-2Q=2Q$$

$$40=4Q$$

$$Q=10$$

To check that the supply curve is increasing, we can compare the slope of MR and MC . If the slope of MC is greater than slope of MR , the second condition is fulfilled.

Second order condition: slope of

Slope of

Thus, the profit maximizing level of output is 10 and the profit maximizing price is obtained by substituting the profit maximizing quantity (10) in the demand function.

Thus, $P = 40 - Q$

$$P = 40 - 10 = 30$$

b) The maximum profit is the level of profit obtained from selling 10 units at \$ 30 each.

$$\Pi = TR - TC$$

But $TR = P \cdot Q$

$$= \$ 30 * 10 = \$ 300$$

$$TC = 50 + Q^2 = 50 + 10^2 = \$ 150$$

The maximum Π is thus $\$ 300 - \$ 150 = \$ 150$.

Remark: there is no unique supply curve for the firm under pure monopoly. This is because there is no one to one correspondence between price and quantity implying that a given quantity could be supplied at different prices and different quantities can be sold at the same price, depending on market demand and marginal revenue.

6.2.6. Long – Run Equilibrium under Monopoly

The monopolist's long run condition is different from the perfectly competitive firms' long run situation in respect of the entry of new firms into an industry. Unlike a competitive firm, the monopolist firm can enjoy a positive profit even in the long run because there are entry barriers that discourage new firms to enter the industry, attracted by the positive profit.

A monopolist maximizes its long run profit when it produces and sells that output level at the falling part of its LAC where $LMC = MR$ ($P > MC$), slope of LMC being greater than the slope of MR at the point of intersection, and the optimal plant size is the one whose SAC curve is tangent to the LAC at the point corresponding to long run equilibrium output.

6.2.7. Social Costs of Monopoly

The question of whether perfect competition or monopoly is more desirable from society's point of view can be explained in terms of price and output in the two types of market. It is believed that welfare of the society is maximized when price equals marginal cost ($P=MC$). In the competitive market equilibrium occurs where price equals marginal cost ($P=MC$). However, under monopoly price exceeds marginal cost ($P>MC$). This implies that competitive market maximizes welfare because $P=MC$; but monopoly imposes a deadweight loss to the society because it causes consumers to buy less output and pay high price. The other disadvantage (Social cost) of monopoly is that it discourages innovations. Monopolist may feel secure and have no incentive to innovate new product (technology) since there are no competitors.

Review questions

1. The MR curve lies below the demand curve in this figure because the:
 - a. demand curve is linear (a straight line).
 - b. demand curve is highly inelastic throughout its full length.
 - c. demand curve is highly elastic throughout its full length.
 - d. gain in revenue from an extra unit of output is less than the price charged for that unit of output.

2. The pure monopolist:

- a. charges the highest price that it could achieve.
- b. earns only a normal profit in the long run.
- c. restricts output to create an insurmountable entry barrier.
- d. restricts output to increase its price and total economic profit.

3. At the monopolist's profit-maximizing output:

- a. price equals marginal revenue.
- b. price equals marginal cost.
- c. price exceeds marginal cost.
- d. profit per unit is maximized.

4. Suppose a monopolist faces a market demand curve given by

$$P = 55 - 2Q \text{ and its total costs of production given by } TC = 100 - 5Q + Q$$

- a. Determine the profit maximizing price and output.
- b. Calculate the dead weight loss from the monopoly power.

6.3. Monopolistic Competition

6.3.1. Definition

Monopolistic competition is a market structure in which a relatively large number of sellers compete to sell a differentiated product. It is a blend of competition and monopoly.

6.3.2. Characteristics of Monopolistic Competition

Monopolistic competition is characterized by:

173. A relatively large number of sellers. The presence of large number of sellers implies that there is:

a. Small market shares Each firm has a comparatively small percentage of the total market and consequently has limited control over market price.

b. No collusion The presence of a relatively large number of firms ensures that collusion by a group of firms to restrict output and set prices is unlikely.

c. Independent action With numerous firms in an industry, there is no feeling of interdependence among them; each firm can determine its own pricing policy without considering the possible reactions of rival firms. A single firm may realize a modest increase in sales by cutting its price, but the effect of that action on competitors' sales will be nearly imperceptible and will probably trigger no response.

2. Differentiated products (often promoted by heavy advertising). There is widespread *non-price competition*, a selling strategy in which one firm tries to distinguish its product or service from all competing products on the basis of attributes like design and workmanship. Product differentiation may be real or imaginary and can be created through advertising. However, the availability of close substitutes severely limits the "monopoly" power of each firm. This means monopolistically competitive industries are much more competitive than they are monopolistic.

3. Easy entry to, and exit from, the industry. Entry into monopolistically competitive industries is relatively easy compared to oligopoly or pure monopoly. Easy of entry limits long-run positive profits and end up with zero profit.

Monopolistic competition share most of the feature of perfect competition and some features of monopoly. The competitive elements result from the large number of firms and the easy entry. The monopoly element results from differentiated (i.e., similar but not identical) products or services. Like the monopolist, the monopolistic competitive firm has some monopoly power so the firm faces a downward sloping demand curve where MR is below the price.

Profit Maximization

The monopolistic competitor faces a demand curve which is negatively sloped (because of product differentiation) but highly elastic (because of the availability of close substitutes). The monopolistic competitor's profit- maximizing or best level of output is the output at which $MR = MC$ ($P > MC$), provided $P > AVC$. At that output, the firm can make a profit, break even, or minimize losses in the short run.

In the long run, firms are either attracted into an industry by short-run profits or leave it if faced with losses until the demand curve facing remaining firms is tangent to its AC curve, and the firm breaks even ($P = AC$). So like a perfect competitor, a monopolistic competitor will earn only a normal profit in the long run or, in other words, will only break even.

Review questions

1. Why does the demand curve of a monopolistic competitor shift down when more firms start production?
2. Which of the following pairs are both "competition like elements" in monopolistic competition?
 - a. Price exceeds MR; standardized product.
 - b. Entry is relatively easy; only a normal profit in the long run.
 - c. Price equals MC at the profit-maximizing output; economic profits are likely in the long run.
 - d. The firms' demand curve is down ward sloping; differentiated products.
3. Suppose a Monopolistically Competitive firm faces a demand function $Q = 160 - P$ and has a Cost Function, $C = 3200 + q^2$ where $MC = 2q$. What quantity and price does the firm charge? Would you expect this firm to face increased competition in the Long Run?

6.4. Oligopoly

Definition

Oligopoly market structure involves only a few sellers of a standardized or differentiated product; so each firm is affected by the decisions of its rivals and must take those decisions into account in determining its own price and output. Oligopoly requires strategic thinking, unlike perfect competition, monopoly, and monopolistic competition.

Characteristics of Oligopoly

The oligopoly market model is developed based on the following major assumptions.

- i) **Very few sellers of the product:** There are very few sellers in the market producing either homogeneous or differentiated products. Pricing and output of a single firm will influence the price and output of the industry. Each firm has large share of the market and it can affect the market activities by altering its output.

ii) Barriers to entry. The same barriers to entry that create pure monopoly also contribute to the creation of oligopoly. Economies of scale are important entry barriers in a number of oligopolistic industries, such as the aircraft, rubber, and copper industries. A closely related barrier is the large expenditure for capital—the cost of obtaining necessary plant and equipment—required for entering certain industries. The ownership and control of raw materials help explain why oligopoly exists in many mining industries, including gold, silver, and copper. In the electronics, chemicals, photographic equipment, office equipment, and pharmaceutical industries, patents have served as entry barriers.

ii) Control over price, but recognized mutual Interdependence:

Because firms are few in oligopolistic industries, each firm is a “price maker”; like the monopolist, it can set its price and output levels to maximize its profit. But unlike the monopolist, which has no rivals, the oligopolist must consider how its rivals will react to any change in its price, output, product characteristics, or advertising. Oligopoly is thus characterized by **strategic behavior** and **mutual interdependence**. By **strategic behavior**, we simply mean self-interested behavior that takes into account the reactions of others. Firms develop and implement price, quality, location, service, and advertising strategies to “grow their business” and expand their profits. But because rivals are few, there is **mutual interdependence**: a situation in which each firm’s profit depends not entirely on its own price and sales strategies but also on those of the other firms. So oligopolistic firms base their decisions on how they think rivals will react. For example; In deciding on its advertising strategy, Burger King will take into consideration how McDonald’s might react.

iii) Presence of monopoly power: There are a few sellers in oligopoly that is why collusion of firms is possible. When they come to an agreement, they will act as monopoly so that high price is charged for the product. They will enjoy more than normal profits by raising the price. In this manner, firms will enjoy monopoly power to some extent.

iv) Homogeneous or differentiated product:

An oligopoly may be either a **homogeneous oligopoly** or a **differentiated oligopoly**, depending on whether the firms in the oligopoly produce standardized or differentiated products. Many industrial products (steel, zinc, copper, aluminum, lead, cement, industrial alcohol) are virtually standardized products that are produced in oligopolies. Alternatively, many consumer goods industries (automobiles, tires, household appliances, electronics equipment, breakfast cereals, cigarettes, and many sporting goods) are differentiated oligopolies. These differentiated oligopolies typically engage in considerable non-price competition supported by heavy advertising.

Types of Oligopoly

Depending on whether there exists some form of cooperation between firms or not, there are two different forms of oligopoly. These are non-collusive and collusive oligopoly. While non-collusive oligopoly signifies the absence of any form of cooperation between firms, collusive oligopoly implies some form of cooperation.

Collusive Oligopoly

The firms in oligopolist markets tend to **collude** to set price for their output. There are two types of collusion of oligopolist firms (a) **Explicit collusion** where oligopolist firms openly collude to determine price and output levels for their product where there are no strong Anti-Trust Laws and (b) **Implicit collusion** where oligopolists agree in more covert ways to fix price and output levels. The latter type of collusion is used in places where Anti-Trust Laws are present and strong. The **collusion of oligopolist firms** in principle lead to the creation of cartels i.e. **cartelization** of oligopolist industries results from their collusive behavior. A cartel is a group of producers (i.e. a producer cartel) or consumers (i.e. consumer cartel) that determines the price and quantity of output to be produced (producer cartel) or consumed (consumer cartel) in a market. For example the Organization of Petroleum Producing Countries or OPEC is a global producer oil cartel.

Profit: While the oligopolist can make profits, break even, or incur losses in the short run, in the long run the firm will leave the industry rather than incur losses. Oligopolists under allocate resources (as $P > MC$) and can earn long-run profits because of restricted entry. Usually they also engage in excessive advertising and product differentiation. The oligopolist usually does not produce at the lowest point on its LAC curve.

The following table provides a summary of the four markets.

Table 6.1. Summary of the major characteristics of the four market structures/models

Characteristics		Market models		
	Pure Competition	Monopolistic Competition	Oligopoly	Pure Monopoly
Number of firms	Large	Many	Few	One or Single
Type of product	Standardized	Differentiated	Standardized or differentiated	Unique no close substitutes
Control over price	None	Some, but with increasing returns to scale; narrow limits	Limited by mutual interdependence; and collusion	Controlled by a single firm

Condition of entry	very easy	relatively easy	significantly barriers/ obstacles/tacitales	Blocked
Non-price competition	None	considerable emphasis on advertising, branding, name, trademarks, emarks	Typically greater than dea l, participation, larl y, width, product differentiation, entiated	Mos tly public relations advertising, marketing
Examples	Agriculture	retail, trade, dresses, shoes	Steele, automation, obsolescence, many households, appliances	Local, utility, ilit ies

Review Questions on Oligopoly

176. What are some of the natural and artificial barriers to entry into oligopolistic industries?

Chapter Seven: Macroeconomics

As we remember, economics is broadly divided into two main branches – microeconomics and macroeconomics. We already covered microeconomics in the previous six chapters. In this part, macroeconomics will be discussed. Here, problems of the macro economy are identified and examined in a reasonable detail. Moreover, national income accounting of the macro economy is introduced and the basic concepts associated to this are clarified, and methods of income accounting are described. Furthermore, among the most pressing macroeconomic problems, unemployment and inflation are separately treated and analyzed.

7.1. Problems of Macroeconomics

Let's first define what macroeconomics is about. Macroeconomics (from Greek prefix *macr(o)* - meaning large + *economics*) is the branch of economics dealing with the structure, performance, behavior and decision making of the *entire economy*. This includes a national, regional or global economy. With this very nature of macroeconomics, macroeconomists develop models that explain the relationship between such factors as *national income, output, consumption, unemployment, inflation, savings, investment, international trade and international finance*. In contrast, microeconomics is primarily focused on the actions of individual agents, such as firms and consumers, and how their behavior determines prices and quantities in specific markets. An important question, however, is why macroeconomics was created as a separate discipline and how it evolved from time immemorial?

Until 1930's there was little need to distinguish | to the Great Depression (Great Economic Crisis) of 1930's classical economists (Economists during 18th century) believed that markets would adjust quickly and direct the economy toward full employment. There was nearly complete unanimity among economists that Even if cyclical problems would happen in the economy they are only temporary or short-lived. They can adjust themselves to equilibrium situation through automatic actions of economic forces. This was the common belief before Great Depression. However, the huge decline in output, prolonged unemployment, and lengthy duration of the Great Depression undermined the classical view and provided the foundation for Keynesian economics. During this period, Keynes started questioning that "in the long run we all are dead why should we wait until the economy collapses". Hence, Keynes successfully challenged classical economists and came up with macroeconomics in 1936. Thus, Great Depression was the very reason for the appearance of the book of Keynes and the rise of interest for macroeconomics.

Having understood why and when macroeconomics was created, what are the main macroeconomic problems countries face with which are supposed to be addressed by macroeconomics? Any economy faces macroeconomic problems such as undesirable situations that exist in the macro economy, largely because one or more of the macroeconomic goals are not satisfactorily attained. The primary macroeconomic problems are unemployment, inflation, maintaining balance of payment and stagnant growth. Macroeconomic theories are, then, designed to explain why these problems emerge and to recommend corrective policies.

Macroeconomic problems arise when the macro economy does not satisfactorily achieve the goals of full employment, price stability, balance of payment and economic growth. Unemployment results when the goal of full employment is not achieved. Inflation exists when the economy falls short of the goal of price stability. These problems are caused by too little or too much demand for gross production. Unemployment results from too little demand and inflation emerges with too much demand. Problem of balance of payment happens when payments to foreigners is more than receipts from foreigners. Stagnant growth means the economy is not adequately attaining the goal of economic growth. Each of these situations is problematic because society is less well off than it would be by reaching the goals.

Issues of growth are generally considered in a long-run framework, which focuses on supply. Business cycles are generally considered in the short-run framework, which focuses on demand. Unemployment and inflation fall within both frameworks

177. Unemployment

Unemployment, as defined by the International Labor Organization, occurs when people are without jobs and they have actively looked for work within the past four weeks. In other words, unemployment is a phase where people who are fit and willing to work do not get work at the prevailing rate of wages after due efforts. The unemployment rate is a measure of the prevalence of unemployment and it is calculated as a percentage by dividing the number of unemployed individuals by all individuals currently in the labor force. i.e.

$$\text{Unemployment rate} = (\text{number of unemployed labor force} / \text{total active labor force}) * 100$$

Unemployment arises when factors of production that are willing and able to produce goods and services are not actively engaged in

production. Unemployment means the economy is not attaining the macroeconomic goal of full employment.

The problem of unemployment is a world-wide reality. The developed countries like the U.S., England, France, Germany, Italy, etc. also suffer from this problem, but it is more pronounced in developing countries. With the passage of time it has become worse. It has become a threat to developing countries' economic well-being and social development. It is one of the major causes of our poverty, backwardness, crimes and frustration among the people. But because of large scale unemployment there is no suitable employment for these people. They are forced to remain idle.

Unemployment is a problem because:

- Less output is produced and thus the economy is less able to address the problem scarcity.
- The owners of unemployed resources receive less income and thus have lower living standards.

Types of unemployment

There are three main types of unemployment. These are: Frictional unemployment, cyclical unemployment and structural unemployment.

- **Frictional unemployment** takes place when *people look for jobs and jobs look for people and the two do not meet*. Discussions of frictional unemployment focus on voluntary decisions to work based on each individual's valuation of one's own work and how that compares to current wage rates plus the time and effort required to find a job. It is temporal unemployment.
- **Cyclical unemployment (seasonal)** is the type of unemployment when *persons look for jobs but there are no jobs*, which can absorb them. Here, it is true that some jobs are dependent up on weather and season.
- **Structural unemployment** occurs due to *mismatch between supply and demand* for workers. It is unemployment as result of a change in the structure of the economy .This is long run unemployment. Structural arguments emphasize causes and solutions related to disruptive technologies and globalization.

2. Inflation

Price instability is a situation in which prices change frequently. It is a situation in which the prices of goods alter daily or even hourly. Price instability is not desirable for the health of any economy. Hence, one of a government's macroeconomic goals is maintaining price stability. Another way to express this is to say that governments desire a low and stable rate of inflation. Then, what is inflation? Is a onetime increase in price considered as inflation? Can we say that inflation is taking place when there is an increase in the price of a particular good or service? In economics, **inflation** is an upward movement in the *average level* of prices. Its opposite is deflation, a downward movement in the average level of prices. The boundary between inflation and deflation is price stability.

Inflation arises when the average price level in the economy consistently and persistently increases. In other words, prices generally rise from month to month and year to year. With inflation the economy is not attaining the stability goal. Inflation is an average increase in prices, with some prices rising more than the average, some raising less, and some even declining. As such, not every member of society is likely to experience exactly the same inflation.

To put it differently, inflation is a *persistent increase* in the average price level in the economy, usually measured through the calculation of a Consumer Price Index (CPI). *What is, then, consumer price index? CPI is the type of price index that is commonly used to measure inflation. CPI weighs each price of the commodity according to their importance to the economy.*

The word "persistent" is of great importance in your understanding of the concept of inflation. A single increase in prices is not called inflation. When inflation occurs, there is a sustained increase in the price level. It is also very important not to confuse inflation with an increase in the price of a *particular good or service*. When the general price level rises, each unit of currency buys fewer goods and services. Consequently, inflation also reflects erosion in the purchasing power of money – a loss of real value in the internal medium of exchange and unit of account in the economy. A chief measure of price inflation is the inflation rate, the annualized percentage change in a general price index (normally the Consumer Price Index) over time.

Inflation's effects on an economy are various and can be simultaneously *positive* and *negative*. Negative effects of inflation include a decrease in the *real value of money/purchasing power of money* and *other monetary items* over time, uncertainty over future inflation may discourage *investment* and *savings*, and high inflation may lead to *shortages of goods* if consumers begin hoarding out of concern that prices will increase

in the future. Positive effects include ensuring central banks can adjust nominal interest rates and encouraging investment in non-monetary capital projects. In general, inflation brings economical, social, political and global harms to the society.

Types and causes of inflation

Inflation can be classified based on different factors such as rate of price change, time, coverage, level of factor of production employment, government interference in the market economy, causes of inflation and etc. However, for this class, we only focus on the causes of inflation. Accordingly, we can broadly classify it to three.

A. Cost push inflation occurs when the costs of production push up the general price level. thus , this is inflation from supply side of the economy .it may be caused by

- Increasing in cost of production such as labor, capital, land etc.
- Imported inflation: the cost of essential imports rise regardless of whether or not they are in short supply. This causes inflation when the price of imported inputs rises , the domestic producers who are using these imported inputs face an increasing cost of production which force them to push up the price of their products and thereby cause inflation.
- Mark –up pricing: many large firms fix their prices on unit cost plus profit basis .This makes price to be sensitive to supply than to demand influences .This means that prices tend to rise automatically with rising costs, whatever the state of economy.

B. Demand pull inflation: it occurs when the aggregate demand exceeds that of aggregate supply .however, in this regard, there is debating among economists. For some group of economists (like classical), inflation happen if aggregate demand is greater than aggregate supply at full employment level of economy. Others (Keynesians) believe that it can be before the economy reaches full employment level. Hence, demand pull inflation may be caused by rising of government expenditure, over private investment, high house hold consumption and others as compare to their contribution to supply side of the economy.

C. Cost push –Demand pull cum inflation (Mixed inflation). In practical world, we may not find single factor that causes of inflation. Thus , inflation can be caused from the two side of the economy i.e. demand push and supply stagnant .Any inflationary phenomenon observed any in the economy may be as result of combination of cost push and demand pull factors

D. Money inflation: money supply excess as compare to output growth may be a cause for inflation to occur in the given economy. Hence, according to Milton Freid Man, inflation is always and every where a monetary phenomenon in the sense that it can be caused by a more rapid increase of quantity of money than in the quantity of out put in the given economy. However, there are many criticisms on the conclusion because it is believed that inflation may be caused by factors other than money supply excess over output.

3. Stagnant Growth

The third problem of macro economy that is stagnant growth arises because the *supply of aggregate production* is not increasing at a desired level or is even declining. An increase in the total production of goods and services is generally needed to keep pace with an increase in the population of society and expectations of a rising living standard. Stagnant growth exists if total production does not keep pace with population growth. This means the macroeconomic goal of economic growth is not attained.

Reasons for stagnant growth can be identified with a closer look at the quantity and quality of the resources used for production.

• *Quantity*: The available quantities of the four factors of production--labor, capital, land, and entrepreneurship--can restrict the growth of production.

The quantity of labor is based on both the overall population and the portion of the population willing and able to work. If one of these decline, then growth is not likely to keep pace with expectations. If, for example, Abebe decides to quit his job and spend his time doing nothing but vegetating on his parents' living room sofa, then the total quantity of labor declines.

• *Quality*: The quality of the four resources can also lead to stagnant growth. The two most noted resource quality influences are technology and education. The lack of technological progress, which could result from allocating fewer resources to scientific research can limit

increases in the quantity of resources. Along a similar line of reasoning, allocating fewer resources to education can also limit resource quality.

4. Problem of Balance of Payment

What does Balance Of Payments - BOP mean? A BOP is a *record of all transactions* made between one particular country and all other countries during a specified period of time. BOP compares the dollar difference of the amount of exports and imports, including all financial exports and imports. A negative balance of payments means that more money is flowing out of the country than coming in, and vice versa.

Balance of payments may be used as an indicator of economic and political stability. For example, if a country has a consistently positive BOP, this could mean that there is significant foreign investment within that country. It may also mean that the country does not export much of its currency. This is just another economic indicator of a country's relative value and, along with all other indicators, should be used with caution. The BOP includes the trade balance, foreign investments and investments by foreigners.

To sum up, the **balance of payments** (BOP) records all of the many financial transactions that are made between consumers, businesses and the government in the country (for Example Ethiopia) with people across the rest of the World. The BOP figures tell us about how much is being spent by Ethiopian consumers and firms on imported goods and services, and how successful Ethiopian firms have been in exporting to other countries and markets. It is an important measure of the relative performance of Ethiopia in the global economy.

When we see BOP of our country, Ethiopia's balance of payments has been significantly affected by weather conditions, terms of trade, and emergency drought relief efforts provided by the international community. The US Central Intelligence Agency (CIA) reports that in 2000 the purchasing power parity of Ethiopia's exports was \$442 million while imports totaled \$1.54 billion resulting in a trade deficit of \$1.098 billion.

The International Monetary Fund (IMF) reports that in 2001 Ethiopia had exports of goods totaling \$433 million and imports totaling \$1.63 billion. The services credit totaled \$523 million and debit \$526 million.

7.2. National Income Accounting

In ordinary business accounting a *firm's performance* can be determined by measuring its profits or loss over a specific period of time. By the same principle, National Income Accounting summarizes a country's economic performance by measuring its aggregate income and output over a given period of time. What is National Income Accounting? National Income Accounting is a *systematic recording of the economic performance of a nation (country)*.

National income accounts (NIAs) are fundamental aggregate statistics in macroeconomic analysis. The ground-breaking development of national income and systems of NIAs was one of the most far-reaching innovations in applied economics in the early twentieth century. NIAs provide a quantitative basis for choosing and assessing economic policies as well as making possible quantitative macroeconomic modeling and analysis. NIAs cannot substitute for policymakers' judgment or allow them to evade policy decisions, but they do provide a basis for the objective statement and assessment of economic policies.

Without national income accounting economic policy would be based on guess work. To avoid the guess work and to assess the health of the economy and to formulate possible policies that can maintain the country, we need National Income Accounting.

7.2.1. Basic Concepts in NIA

A variety of **measures of national income and output** are used in economics to estimate total economic activity in a country or region, including gross domestic product (**GDP**), gross national product (**GNP**), and net national income (**NNI**). All are especially concerned with counting the total amount of goods and services produced within some "boundary". The boundary is usually defined by geography or citizenship, and may also restrict the goods and services that are counted. For instance, some measures count only goods and services that are exchanged for money, excluding bartered goods (exchange of goods for goods), while other measures may attempt to include bartered goods by *imputing* monetary values to them.

National Income

National income is a measure of the value of goods and services produced by the residents of an economy in a given period of time, usually a quarter or a year. National income can be real or nominal. Nominal national income refers to the current year production of goods and services valued at current year prices. Real national income refers to the current year production of goods and services valued at base year prices (the normal year where there is no inflation or deflation).

In estimating national income, only productive activities are included in the computation of national income. In addition, only the values of goods and services produced in the current year are included in the computation of national income. Hence, gains from resale are excluded but the services provided by the agents are counted. Similarly, transfer payments are excluded as there is income received but no good or service produced in return. However, not all goods and services from productive activities enter into market transactions. Hence, imputations are made for these non-marketed but productive activities. Imputed rental for owner-occupied housing is a typical example. Thus, national income refers to the market value or imputed value of additional goods and services produced and services performed in the current period.

GDP, GNP, NDP and NNP.

National incomes in many countries are either in Gross Domestic Product (GDP) or Gross National Product (GNP). Gross Domestic Product (GDP) refers to the *total value of goods and services produced within the geographical boundary of a country before the deduction of capital consumption*. Net Domestic Product (NDP) refers to the total value of goods and services produced within the geographical boundary of a country *after the deduction of capital consumption (depreciation of capital goods)*.

Gross National Product (GNP) refers to the *total value of goods and services produced by productive factors owned by residents of the country both inside and outside of the country before the deduction of capital consumption*. Net National Product (NNP) refers to the total value of goods and services produced by productive factors owned by residents of the country both inside and outside of the country *after the deduction of capital consumption (depreciation of capital goods)*.

Relationship between GDP and GNP

$$\text{GNP} = \text{GDP} + \text{NPIFA} \text{ (Net Property Income from Abroad)}$$

Net Property Income from abroad refers to the difference between income that comes from abroad and income that goes to abroad.

7.2.2. Approaches to Measuring GNP: Product, Expenditure and Income

There are 3 approaches to measure national income i.e. output approach, income approach and expenditure approach. Theoretically, the national income calculated from the 3 approaches is the same, i.e.,

$$\text{Expenditure Approach} = \text{Output Approach} = \text{Income Approach} \text{ Or}$$

$$\text{Expenditure on goods} = \text{Value of goods and services produced} = \text{Income Earned by services of factor owners}$$

a. Output Approach

Output approach measures national income by adding the total value of the *final goods and services produced in the year* or by adding the value added by each sector of the economy.

Value Added

Value added refers to the difference between the value of gross output of all goods and services produced in a given period and the value of *intermediate inputs* used in the production process during the same period. In the value added approach, national income is calculated as the value of sales of goods - purchase of intermediate goods to produce the goods sold.

Double counting or multiple counting of the value of wheat, for example, at each production stage is avoided when the *value added or final value* is used in the computation as the value of wheat is counted only once in the value added method or the final value method. In the value added approach, national income is difference between the value of sales of goods - purchase of intermediate goods to produce the goods sold.

b. Income Approach

Income approach measures national income by adding the income earned by the factor owners that are residents of the country, undistributed company profits and government income from economic participation. It excludes transfer payments and stock appreciation because transfer payments and stock appreciation are not due to goods and services performed.

Computation of National Income by Income Approach includes:

Income from Employment

Income from Self-employment

Gross trading profits of companies

Gross trading surplus of public corporations

Gross trading surplus of general government enterprises

Rent, Interest and Dividend, etc

In other words, using the Income Approach, GNP is calculated by adding up the factor incomes of the factors of production in the society. These include

National Income (NY) + Indirect Business Taxes (IBT) + Capital Consumption Allowance and Depreciation (CCA) + Net Factor Payments to the rest of the world (NFP)

In this approach,

NY (National Income) = Employee compensation + Corporate profits + Proprietor's Income + Rental income + Net Interest

NFP = Payments of factor income to the rest of the world minus the receipt of factor income from the rest of the world.

Thus,

GNP - NFP = GDP (GROSS DOMESTIC PRODUCT)

GNP - CCA = NNP (NET NATIONAL PRODUCT)

NNP - IBT = NY (NATIONAL INCOME)

Table 1: National income using income approach

Compensation of employees	4,008.3
Proprietors' income with adjustments	450.9
Rental income of persons with adjustments	116.6
Corporate profits with adjustments	526.5
Profits before tax	528.2
Profits tax liability	195.3
Dividends	211.0
Undistributed profits	121.9
Inventory valuation and capital consumption adjustments	-13.3
Net interest	392.8
NATIONAL INCOME(NI)	5495.1

Note: Data are billions of nominal dollars

c. Expenditure Approach

Expenditure Approach measures national income by adding the private consumption expenditure, government consumption expenditure, gross fixed capital formation i.e. investment expenditure, increase in physical stocks and net exports of goods and services i.e. the difference between exports and imports.

It only includes expenditure on goods and services to satisfy the needs of final buyers. It excludes expenditure on intermediate goods and services. Moreover, resale of consumer and capital goods are excluded because the expenditures are on these resale goods, not goods produced in the current period and hence expenditures on resale goods are not counted.

Private Consumption Expenditure (consumption)

Private consumption expenditure refers to the final purchases of goods and services by households. It includes expenditure on single use consumption or non-durable goods e.g. food, durable goods such as TV, household utensils, and services including hairdressing services and medical services.

Household purchases of *new houses are treated as investment* expenditure and hence residential investments are not included in private consumption expenditure. Instead residential investments are included in investment expenditure. Resale of consumer durables e.g. second hand TV are excluded as the expenditures are on second hand TV, not TV produced in the current year and hence expenditures on second hand TV are not included in the private consumption expenditure.

Government Consumption Expenditure / Government Expenditure

Government consumption expenditure refers to the cost of running the various *government departments and public non-profit organizations* to provide goods and services for the public. It excludes the expenditure by government on grants, interest subsidies, transfer payments, loans and repayments.

Gross Domestic Fixed Capital Formation / Investment Expenditure

Investment expenditure refers to the expenditure on equipments and machinery, residential and non-residential construction, and changes in inventories. An increase in inventories is treated as an investment and a fall in inventories is treated as dis-investment. Resales of capital goods are excluded from investment expenditure.

Gross Investment = Net Investment + Replacement Investment

Gross investment refers to the new capital goods produced that can be used to the capital stock or to replace the existing worn-out capital goods. Net Investment refers to the new capital goods that are added to the capital stock. Replacement investment refers to the new capital goods that are used to replace the existing worn-out capital goods.

Net Exports of Goods and Services

Net exports of goods and services refer to the difference between the exports of goods and services and imports of goods and services. Exports of goods and services refer to goods and services that are produced in the country but they are sold to foreigners for their consumption. Imports of goods and services refer to goods and services that are produced by other countries but they are consumed within the country.

Computation of National Income by Expenditure Approach

National Income =Private Consumption Expenditure + Government Consumption Expenditure + Gross Domestic Fixed Capital Formation / Investment Expenditure+ Exports of Goods and Services

In other words, GDP is the total spending on all final goods and services (Consumption goods and services (C) + Gross Investments (I) + Government Purchases (G) + (Exports (X) - Imports (M))

$$GDP = C + I + G + (X - M)$$

Table 2: GDP computation based on Expenditure approach

Component of GDP	(billions of dollars)	(dollars)
House hold Consumption		6,757.
Nondurable goods		2,010.
Durable goods		820.
Services		3,927.0
Investment		1,832.7
Nonresidential fixed investment		1,362.2
Residential fixed investment		416.0
Inventory investment		54.5
Government Purchases		1,743.7
Federal		595.2
Defense		377.0
Nondefense		218.2
State and local		1,148.6
Net Exports		□ 370.7
Exports		1,097.3
Imports	1,468	
Gross Domestic Product		9,963.

Source: U.S. Department of Commerce. GDP and the Components of Expenditure: 2000

7.2.3. Nominal and Real GDP

Nominal GDP measures the value of output during a given year using the prices prevailing during that year. Over time, the general level of prices rises due to inflation, leading to an increase in nominal GDP even if the volume of goods and services produced is unchanged.

Real GDP measures the value of output in two or more different years by valuing the goods and services adjusted for inflation. For example, if both the "nominal GDP" and price level doubled between 1995 and 2005, the "real GDP" would remain the same. For year over year GDP growth, "real GDP" is usually used as it gives a more accurate view of the economy.

Relation between Real GDP and Nominal GDP

Real GDP is calculated using constant prices (base year prices) whereas nominal GDP uses current prices. The difference between the nominal GDP and real GDP is due to the inflation rate in market.

Example: Assume our simplistic economy only produces teff and maize. The price for teff is 800 Birr in 2000, whereas the price for maize is 500 Birr. Same year we produce 100 tons of teff and 50 tons of maize. In 2003, because of the inflation the price for teff goes up to 1000 Birr, whereas the price for maize is 600 Birr at the same production levels.

The nominal GDP in 2000 is 105000 ($800*100 + 500*50$) and the nominal GDP in 2003 is 130000($1000*100+ 600*50$)

Real GDP at 2000 is 105000 ($800*100 + 500*50$) and Real GDP at 2003 is 105000 ($800*100 + 500*50$). Hence, real GDP did not change, because real GDP only changes with the changing production level and therefore is a better size measure for an economy.

Review questions

- i. Why macroeconomics is needed?
- ii. What are the main macroeconomic problems in developing countries?
- iii. What are the problems in using national income accounting?
- iv. What are the similarities and differences between GDP and GNP?
- v. Can you state the possible problems in measuring GDP correctly?
- vi. From the following table , compute
- d. Gross National product (GDP)
- e. Gross Domestic Product(GNP)

f. National Income (NI)

TABLE: National income accounting data

Personal Consumption Expenditure	4,698.7
Gross private domestic investment	1,014.4
Producer's durable equipment and nonresidential structures	667.2
Residential structures	287.7
Changes in business inventories	59.5
Net exports of goods and services	-96.4
	722.0
• Exports	818.4
• Imports	1,314.7
Government purchases of goods and services	516.3
	798.4
• Federal	-9.0
• State and local	818.8
net receipts of factor income from rest of world	608.6
Consumption of fixed capital	
Indirect business taxes, business	
transfer payments, statistical discrepancy, misc.	