Wollo University

College of Agriculture

Department of Animal Science

Introduction to Agri-Economics (Econo 4151)

A Teaching Material

Instructor: ENDRIE M. (MSc)

January, 2020

Wollo

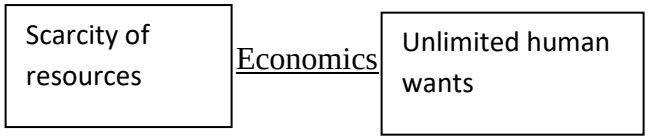
CHAPTER ONE
Nature, Basic Concepts and Methodology of Economics

INTRODUCTION

Most of our public issues are closely related to economic matters. For instance, people may be interested to know the cause of price instability, unemployment and output fluctuation in the economy. In an economy, when there is persistent rise in general price, there will be inflation. When general prices sharply decline, deflation will occur. Both inflation and deflation are not desired or needed for any economy. Hence, the economy should solve the problem through different policies. Citizens should understand the economic problems that affect their standard of living. Therefore, the knowledge of the basic principles of economics is vital.

1.1. Definition of economics

Basically, there is no consensus among economists about a precise definition of economics. There are many definitions of economics, each trying to encapsulate the "fundamentals of the subject."
In principle there is no universally accepted definition of what constitutes the term **"economics."** But for our purpose we can introduce one of various definitions of economics.
Economics is a social science that studies how societies use the limited/scarce resources to produce goods and services that satisfy unlimited human wants and desires.



Human beings want food, cloth, shelter and other variety of goods and services for their survival. These human wants are unlimited and increase from time to time; however economic resources are scarce by their nature.
Economics deals with how a particular society solves the economic problems. For instance:

Problems	**Solutions**
What to produce?	A. Traditional-growing the same crop using the same method
How much to produce?	B. Technical-selection of the highest possible alternatives
How to produce?	C. Economic-produce based on least cost combination methods
Where & when to produce?	
For whom to produce?	

Economic problems exist, whenever resources are scarce. Economic theory tries to narrow the gap between the available resources and unlimited human wants.

Economic resource is a thing or service used to produce goods or services that can satisfy human wants. **Goods** are tangible things that can help a consumer to satisfy his/her wants, for instance, car, pencils, table, etc. On the other hand, **services** are intangible things from which the

consumer can derive satisfaction, for instance, medical service, legal service and defense service. The need to balance unlimited wants with limited resources has raised the question of efficient utilization of scarce resources. In other words, economics is a study of how people choose scarce productive resources to produce different commodities in an economy.

An Economy is defined as a mechanism or structure that organizes and allocates scarce resources for maximization of production of goods and services desired by the society.

Economics is doing the best with what we have. Hence, it addresses three issues

- ❖ Scarcity
- ❖ Alternative uses &
- ❖ Objective of profit maximization

Therefore, the basic aim of economics is the efficient use of the scarce resources.

Productive resources are resources which are used for further production. These are:

a) Land
b) Labor
c) Capital
d) Management
e) Technology & marketing

Features of Economics

Economists agree that "each theme" is an essential feature of economics of the subject economics (Agreed by economists are)

A) Scarcity-it is an important feature of economics. If there were "no scarcity", there would be no need to study economics

B) Choice-economics is a science of choice. Scarcity & choice go together. Individuals, businesses, & societies must choose among alternatives.

C) Specialization-economics studies how participants in the economy specialize in tasks to which they are best suited.

D) Exchange-it is another feature/theme of economics which is a complement of specialization. With no exchange, specialization would be of no benefit- because individuals could not trade the goods in which they specialize for those that other individuals produce.

For example, workers exchange specialized labor (building, electricity, medical, etc) for money and exchange money for a variety of goods.

1.2. The Nature of Economics (Scope of Economics)

Economic theory is divided into two parts: microeconomic theory and macroeconomic theory. Microeconomic theory considers economic reasoning from the view point of individuals and firms and builds up from there to an analysis of the whole economy. **Microeconomics** is defined as the study of individual choice, and how that choice is influenced by economic forces. Microeconomics studies such things as the pricing policies of firms, households' decisions on what to buy, and how markets allocate resources among alternative ends. Our discussion of opportunity cost was based on microeconomic theory.

As one builds up from **microeconomic** analysis to an analysis of the entire economy, everything gets rather complicated. Many economists try to uncomplicated matters by taking a different approach-a macroeconomic approach-first looking at the aggregate, or whole, and then breaking it down into components. **Macroeconomics** is defined as the study of the economy as a whole. It considers the problems of inflation, unemployment, business cycles, and growth.

Macroeconomics focuses on aggregate relationships such as how household consumption is related to income and how government policies can affect growth. A micro approach would analyze a person by looking first at each individual cell and then building up. A macro approach would start with the person and then go on to his or her components-arms, legs, fingernails, feelings and so on. Put simply, microeconomics analyzes from the parts to the whole, macroeconomics analyzes from the whole to the parts.

Microeconomics and macroeconomics are very much interrelated. This means, microeconomics and macroeconomics can never be separated from one another because **macroeconomics** is the aggregate of micro level economic activities. For instance, the total income (a macroeconomic issue) is the sum of income of individuals in the economy (a microeconomic issue). Therefore, one can't be studied in isolation from the other.
Economics can also be analyzed as positive and normative economic analysis. Let's examine each of them as follows:

1.3. Methods and Levels of Economic Analysis

The fundamental objective of economics, like any science, is the establishment of valid generalizations about certain aspects of human behavior. These generalizations are known as **theories (**like the law of supply and demand). Theory is a simplified picture of reality that tells us how one thing influences another thing. Why is the establishment of valid generalization (cause and effect relationship) about economic phenomena so important? Because it better enables us to make predictions about the phenomena. Prediction is important because it gives us the possibility to control the outcome. That is, in our study of economics, we need to know what makes people take economic decisions. That knowledge will enable us to predict the consequences of people's behavior. The prediction, in turn, enables us to formulate policies to change the outcome, if the change proves appropriate. For example, the law of demand is a generalization about consumers' behavior. It states that there is an inverse relationship between the price of a good and quantity purchased. This theory will enable the seller to predict the outcome of the change in the price of the good. The prediction will help the seller to design pricing policy which maximizes his sales revenue.

Economic theories are often based on simplifying assumptions in order to reach conclusions. One simplifying assumption is **"Ceteris Paribus"**. This is a Latin phrase which means "If other things remain constant". This assumption is made because there are many economic factors in an actual situation that make the analysis difficult. It must be noted that some of the other things that are assumed to be constant may vary and hence affect the model's result. Economic theory provides the basis for economic analysis. Economic analysis makes use of logical reasoning. Different methods of economic analysis to be applied by different school of thought are:

1.3.1. Positive and Normative Economic Analysis

Positive economics is part of economic science, which concerns itself with statements that are capable of being verified by investigating facts. It deals with how the economic problems facing a society are actually solved. Positive economic analysis is a statement which provides solution for "what is", "what was" or "what will be" rather than "what should be" or "what ought to be".

Positive economics determines who gains and who losses as a result of a change in the economic conditions or policies through verification of facts. It is important to forecast the impact of the changes in the economic conditions or policies on observable items like production, cost, sales, prices, and personal incomes, etc. Suppose the firm pollutes the air in the process of producing its output. If we study how much additional cleaning cost is imposed on the community by this pollution, then we are dealing with positive economics. Statements like "what is the inflation rate this year?", "what percent of the labor force is unemployed?" and "how many people earn less than 1000 Birr per month?" are questions that can be answered by reference to facts.

Normative Economics is the study of what the goals of the economy should be. Normative Economics asks such questions as: what should the distribution of income be? What should tax policy be designed to achieve? In discussing such questions, economists must carefully delineate whose goals they are discussing. One cannot simply assume that one's own goals for society are society's goals.

Normative Economics can be argued about but it can never be settled by investigating facts unlike positive economics. Hence, there is no simple right or wrong answer to normative economics. The normative approach makes recommendations regarding "what ought to be". Statements like "families of six children with incomes below 6000 Birr per year should be exempted from income taxes" and "tax revenues ought to be used to provide incomes to the poor" are some of examples for normative economic analysis. They are questions involving answers based on somebody's value judgment. Whenever words such as "ought" or "should" appear in a sentence, there is strong chance that you are dealing with a normative statement.

1.3.2. Induction and Deduction Methods

The two kinds of methods of analysis which have been adopted at different stages of economic science by different schools of thought are deductive and inductive methods.

The deductive method is also known as analytical method. It is a logical reasoning to explain specific events on the basis of the already established theory. Thus, in deductive method reasoning goes from the general to the particular (from theory to facts). Major steps in the deductive approach:

a) Identification of the problem for analysis,
b) Specification of the assumption,
c) Formulating hypotheses on the basis of assumption, and
d) Testing validity of the hypotheses.

The inductive method is a logical way of reasoning to arrive at a valid general statement, on the basis of valid specific facts. For example, we might observe that employment tends to rise in the summer and fall in the winter. Having observed this regular pattern, we might produce a theory to explain it which could then be tested against the facts in the normal way. In inductive method reasoning proceeds from the particular to the general (from facts to theory). Inductive method involves the following steps:

a) Selecting problem for analysis,

b) Collection, classification, and analysis ofw data by using appropriate statistical techniques in order to establish the relationship between variables, and
c) To find out the reasons for the relationship established through statistical analysis.

While deductive method is a descending process, inductive method is an ascending process. It is now widely accepted that deductive and inductive methods are complementary. A combination of the two methods yields a more reliable economic theory.

1.4. Fundamentals of Economics (Scarcity, Choice, Decision Making, Production possibility Curve, etc)

Natural resources, the supply of workers, managers, equipment, time and so on are **scarce**. That means that in our economy there is a problem of scarcity-the goods available are too few to satisfy individuals' desires. Society can have more of one thing only by reducing the other. This means, society must give up something to obtain a scarce good. The opposite of a scarce good is a free good. An item is a **free good** if it can be obtained free or with no effort. Since we make no choice for free good, it has a zero opportunity cost. Air is an example of a free good because we don't need to give up anything to obtain air. The degree of scarcity is constantly changing. The quantity of goods, services, and useable resources depends on technology and human action, which underlie production. Individuals' imagination, innovativeness, and willingness to do what needs to be done can greatly increase available goods and resources. However, scarcity would not be entirely eliminated since new wants are constantly developing. Scarcity forces people to make choice and choice involves sacrifice. When scarcity exists we must sacrifice something to obtain more of any scarce good or service. Thus, resources can never be enough to meet all society's desire. The more resources are used for one purpose, the less will be available for other purpose. Hence, societies make choices to use scarce resources to satisfy their unlimited wants. When societies make a choice, they forgo something in return, which is an opportunity cost. This means, scarcity implies choice and choice implies sacrifice of the next best alternative, which is called an **opportunity cost**.

1.4.1. The production possibilities curve (PPC)

The problem of scarcity can be shown with the aid of a simple model whose purpose is to examine the relationship between the production of goods and services and the availability and use of resources. At any point in time, given available resources, their quality and current technology, a country can produce only a certain amount of goods and services with its scarce resources. Moreover, the country cannot produce more of one good without giving up (reducing) other good's production. In economics, we represent this limitation on a country's productive potential by the production possibility curve (PPC).

Hence, PPC is a curve that shows the maximum possible output of the combinations of two goods (or broad classes of goods) that can be produced with available resources and given technology at a given period of time when all economic resources are fully (efficiently) utilized. PPC can show the maximum possible output of one good that can be produced with available

resources given the output of the other good over a period. To draw the PPC, make the following assumptions:

a) The quantity and quality of economic resources available for use during the year are fixed.
b) To simplify the illustration, assume that there are two broad classes of output to be produced with available economic resources over the year.
c) Specialization. Some inputs are better adapted to the production of one good than to the production of the other.
d) Technology does not change over the year (fixed technology).
e) Economic efficiency.

Economics is the science that is concerned with the problem of using scarce resources so as to attain the greatest or the maximum fulfillment of society's unlimited wants. If our wants are unlimited and resources are scarce, then we can't satisfy all society's want. So economics is concerned with "doing the best with what we have" to achieve the greatest possible satisfaction of society's material wants. Thus, economics is a science of efficiency. This means efficiency in the use of scarce resources. Efficiency means absence of waste or using the economic resources as efficiently as possible to satisfy people's needs and desires. Therefore, efficiency implies the production of any goods and services in the least cost manner. When we produce at the least cost, we are using the smallest amount of resources and therefore making available the resources to produce other goods and services. The conditions to achieve efficiency are full employment and full production.

1.4.2. Economic Efficiency

1.4.2.1. Full employment

When all available resources are employed and no workers are involuntarily out of work i.e. if the economy provide employment for all who are willing and able to work and no capital equipment or arable land is idle then we can say there is full employment.

1.4.2.2. Full production

The employment of all available resources is not enough to achieve efficiently because full production must be also realized. Full production means that all employed resources should be used so that they provide the maximum possible satisfaction. Full production implies two kinds of efficiency: allocative efficiency and productive efficiency. Allocative efficiency means that resources are being devoted to the combination of goods and services that are most wanted by the society. Productive efficiency will be realized when the desired goods and services are produced in the least costly ways. In short, allocative efficiency means "producing the right good" while productive efficiency means "producing in the right way". To achieve economic efficiency both full employment and full production should be realized.

Assume we have a hypothetical economy that produces only two types of products, teff and wheat. Let us consider the different combinations of the two products that the economy can produce.

Table 1.1. Production possibilities of teff and wheat

	Production alternative				
Types of product	**A**	**B**	**C**	**D**	**E**
Teff (in million tons)	0	1	2	3	4
wheat (in million tons)	10	9	7	4	0

If all the resources are used in the production of wheat, the maximum amount of wheat (10 million tons) will be produced. Similarly, if all the resources are used in the production of teff, the maximum amount of it (4 million tons) will be produced, with no output of wheat. These are two extreme possibilities. In between, various combinations of wheat and teff can be produced. Since resources are limited in supply and fully employed, any increase in the production of teff will necessitate the shifting of resources away from the production of wheat and vice versa.

Figure 1.9 is a graphic presentation of the above example. The horizontal axis measures teff and the vertical axis measures wheat. The five production alternatives are plotted. By joining these points with a smooth line we obtain the production possibility curve. The curve shows that if the society wants to have more teff, it must have less of wheat. Each point on the production possibility curve represents some maximum output of the products. With fixed inputs and given technology, we cannot go beyond the production possibility curve. For example, point **F** is unattainable. If there are idle resources (unemployment), or if employed resources are used inefficiently (underemployment), society is not producing on the production possibility frontier, but rather somewhere inside it. With unemployment or underemployment, society produces less than each alternative shown in Table 1.1. Graphically, this situation can be shown by a point inside the original production possibility curve, like point **G** in Figure 1.9.

The production possibility curve helps us see what is meant by efficiency. At point **G**, in Figure 1.9, all points are used to make 4tons of wheat and 2 tons of teff. This is inefficient since there is a way to obtain more of one without giving up any of the other, that is, to obtain 7 tons of wheat and 2 tons of teff (point **C**) or 3 tons of teff and 4 tons of wheat (point **D**). All points inside the production possibility curve are inefficient.

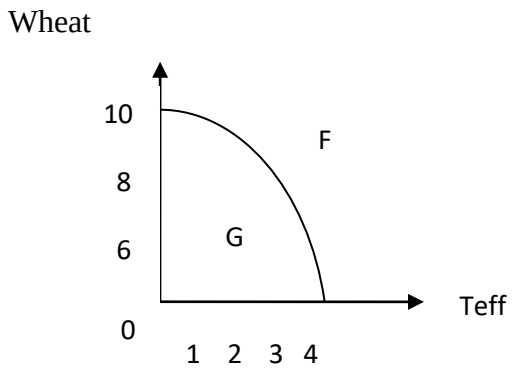
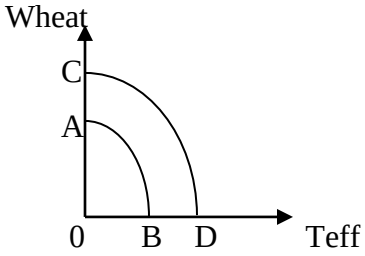
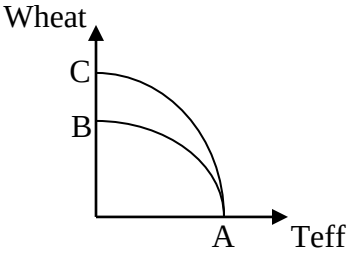


Fig 1.9. Production possibility curve

A technological change that improves production techniques will shift the production possibility curve outward, as shown in Fig 1.10 (**a** and **b**). How the curve shifts outward depends on how technology improves. For example, if we become more efficient in the production of both teff and wheat, the curve will shift out as in (**a**). If we become more efficient in producing wheat, but not in producing teff, then the curve will shift as in (**b**).



a) Neutral technological change



b) Biased technological change

1.4.3. Opportunity cost

The opportunity cost of undertaking an activity is the benefit forgone by undertaking that activity. The benefit foregone is the benefit that you might have gained from choosing the next-best alternative. To obtain the benefit of something, you must give up (forgo) something else-namely, the next-best alternative. All activities that have a next-best alternative have an opportunity cost. For instance, if you are considering the choice between studding for an hour and swimming for an hour, then the opportunity cost of studying is the hour of swimming you sacrifice when you choose to study. Therefore, opportunity cost is a measure of everything you sacrifice to attain a given objective.

We have emphasized that resources are limited relative to the unlimited wants, which these resources can be used to satisfy. As a result, society must make choice among alternatives. As the above table and Fig_ illustrate, initially for producing one million tons of teff, one million tons of wheat have to be given up (alternative **A** to **B**). When the total output of teff is to be increased from one to two, two amounts of wheat have to be sacrificed (alternative **B** to **C**) and so on. The amount of other products which must be foregone (given up) to obtain a unit of any given product is called the opportunity cost of that good.
In equation form,

$$\text{Opportunity cost} = \frac{\textit{The amount sacrificed of one good}}{\textit{The amount obtained of the other good}}$$

In our case, the amount of wheat which must be sacrificed to get another unit of teff is the opportunity cost of teff. For example, when we go from alternative **B** to alternative **C**, the opportunity cost of teff equals 2 ($|\frac{7-9}{2-1}|$).The opportunity cost of wheat, in moving from **C** to **B** equals 0.5 ($|\frac{1-2}{9-7}|$).In moving from alternative **A** to alternative **E**, the cost of wheat involved in getting each additional unit of teff increases. That is when the economy moves from **A** towards **E**, it must give up successively larger amounts of wheat (**1, 2, 3, 4**) to acquire equal increments

of teff (**1, 1, 1, 1,).** This is reflected in the shape of the production possibility curve, particularly the concavity (outward bowed) of the curve illustrates the law of increasing costs.

What is the economic rationale for the law of increasing costs? There are three explanations to this phenomenon:

i) There are diminishing returns. To produce more of teff society needs more and more resources. But, one or more of these resources may be relatively limited. That is, society may begin to run out of one or more of the economic resources. If units of a resource are added to a fixed proportion of other resources, eventually marginal output will decline.
ii) Diseconomies of scale. By shifting from wheat to teff, firms making teff will grow so large that diseconomies of scale will eventually set in.
iii) Factor suitability. When resources are first transferred to produce teff from wheat, those resources which are most productive of teff and least productive of wheat are transferred. This means that resources are not equally adaptable to alternative uses, and hence the lack of perfect flexibility on the part of economic resources and the resulting increase in the sacrifice of one good that must be made in the acquisition of more and more units of another good is the other reason for the law of increasing costs.

If the opportunity costs were constant rather than increasing, the production possibility curve would be a straight line. Finally, since free goods are abundant in supply nothing is sacrificed for their consumption. Thus, free goods have zero opportunity cost.
The above discussion confirms the fact that, with the given assumptions, the production possibility curve cannot expand (shift to the right). However, if the production possibility curve shifts to the right, there is economic growth. Economic growth is the expansion in production possibility that results from the following sources:
a) Increased quantities of economic resources,
b) Improved quality of economic resources, and

c) Advances in technology.

1.5. The Foundation of Economics

Human wants are practically unlimited, while the resources available to produce goods, which satisfy wants, are limited. The main problem in economics is allocating resources in a way, which satisfies as many of these wants as possible.
The following are facts that provide foundations for the field of economics:

- The economic resources are limited or scarce in supply and they have alternative uses.
- Human or society's material wants are unlimited.

The need to balance the more or less unlimited ends with limited means has given rise to the question of efficient utilization of scarce means. Thus, the imbalance between unlimited ends and scarce means provides a foundation for the field of economics. By ‘ends’ we mean wants for goods and services. ‘Means’ refers to resources.

The four broad categories of resources-land, capital, labor and entrepreneurial ability are sometimes known as factors of production. One important characteristic of economic resources is that they can be put to alternative uses. If resources are specific and can be utilized only for

one purpose, the question of choice in the allocation of resources will not arise. However, most economic resources have alternative uses and if we use them to produce one thing, we must forego (give up) the output of some other things. Scarcity forces people to make choice (economic decision) and choice involves sacrifice. Society, therefore, must choose the best way of using scarce resources so as to achieve the maximum level of economic welfare. The problems of choice are essentially problems of allocation. There are three fundamental choices to be made.

a) What to produce. This problem concerns the composition of total output. It implies that the society must determine, by some method, what kinds of goods and services should be produced and how much in each category.

b) How to produce. This entails decisions on the appropriate combinations of different resources for producing a good. It refers to the methods of production to be adopted. Most goods can be produced in a variety of methods. A basic distinction is between capital-intensive production and labor-intensive production. Capital-intensive uses large amounts of capital relative to labor while the labor-intensive uses large amounts of labor relative to capital. These decisions clearly affect the efficiency with which we use our resources and the greater the efficiency, the greater the output that can be produced from any given quantity of inputs.

c) For whom to produce. This is sometimes referred to as the problem of distribution. The distribution of income is as important as its generation. There is no society where all members enjoy the fruits of production equally. The economic system has to determine the relative sizes of the shares going to each household. These three questions of what, how and for whom are common to all societies regardless of their level of development.

CHAPTER TWO
THEORY OF DEMAND, SUPPLY AND UTILITY THEORY

2.1. Theory of Demand

People want lots of things; they "demand" much less than they want because demand means a willingness and ability to pay. Unless you are willing and able to pay for it you may want it, but you don't demand it. This point is an important aspect of demand: The quantity you demand at a low price differs from the quantity you demand at a high price. Specifically, the quantity you demand varies inversely-in the opposite direction-with price.

Prices are the tool by which the market coordinates individuals' desires and limits how much people are willing to buy-how much they demand. When goods became scarce, the market reduces the quantity of those scarce goods people demand; as their prices go up, people buy fewer goods. As goods become abundant, their prices go down, and people want more of them. The invisible hand-the price mechanism-sees to it that what people demand (do what's necessary to get) matches what's available. In doing so, the invisible hand coordinates individuals' demands.

2.1.1. The law of demand and determinants of demand

The **law of demand** states that the quantity of a good demanded is inversely related to the good's price. This means:

- Quantity demanded rises as price falls, other things constant. Or alternatively:
- Quantity demanded falls as price rises, other things constant.

The demand curve

A demand curve is the graphic representation of the relationship between price and quantity demanded. Figure 2.5 shows a demand curve.

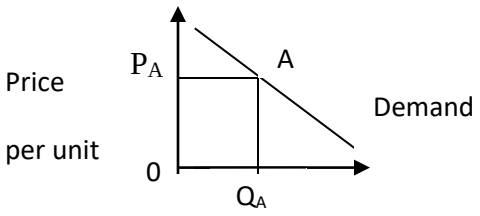


Figure 2.1. A sample demand curve

As you can see, in graphical terms, the law of demand states that as the price goes up, the quantity demanded goes down, other things constant. But what does "other things constant" mean? Say that over a period of two years, both the price of cars and the number of cars purchased rise. That seems to violate the law of demand, since the number of cars purchased should have fallen in response to the rise in price. Looking at the data more closely, however, we see that a third factor has also changed: Individuals' income has increased. As income increases, people buy more cars, increasing the demand for cars. That is why the qualifying phrase "other things constant" is an important part of the law of demand. The other things that are held constant include individuals' tastes, prices of other goods, and even the weather.

2.1.2. Shifts in demand versus movements along a demand curve

To distinguish between the effects of price and the effects of other factors on how much of a good is demanded, economists have developed the following precise terminology. The first distinction to make is between demand and quantity demanded.

- **Demand** refers to a schedule of quantities of a good that will be bought per unit of time at various prices, other things constant.
- **Quantity demanded** refers to a specific amount that will be demanded per unit of time at a specific price, other things constant.

In graphical terms, the term demand refers to the entire demand curve. Demand tells how much of a good will be bought at various prices. Quantity demanded tells how much of a good will be bought at a specific price; it refers to a point on a demand curve, such as point **A** in Figure 3.1.

This terminology allows us to distinguish between changes in quantity demanded and shifts in demand. A change in the quantity demanded refers to the effect of a price change on the quantity demanded. It refers to a movement along a demand curve-the graphical representation of the effect of a change in price on the quantity demanded. A shift in demand refers to the effect of anything other than price on demand.

Shift Factors of Demand

Shift factors of demand are factors that cause shifts in the demand curve. A change in anything that affects demand besides price causes a shift of the entire demand curve.
Important shift factors of demand include:
1. Society's income.
2. The prices of other goods.
3. Tastes.
4. Expectations.
5. Taxes on and subsidies to consumers.

Income: A rise in income increases the demand for goods. For most goods this is true. As individuals' income rises, they can afford more of the goods they want, such as computers, cars or clothing. For other goods, called inferior goods, an increase in income reduces demand. An example is urban mass transit. A person whose income has risen tends to stop riding the bus to work because he/she can afford to buy a car and rent a parking space.

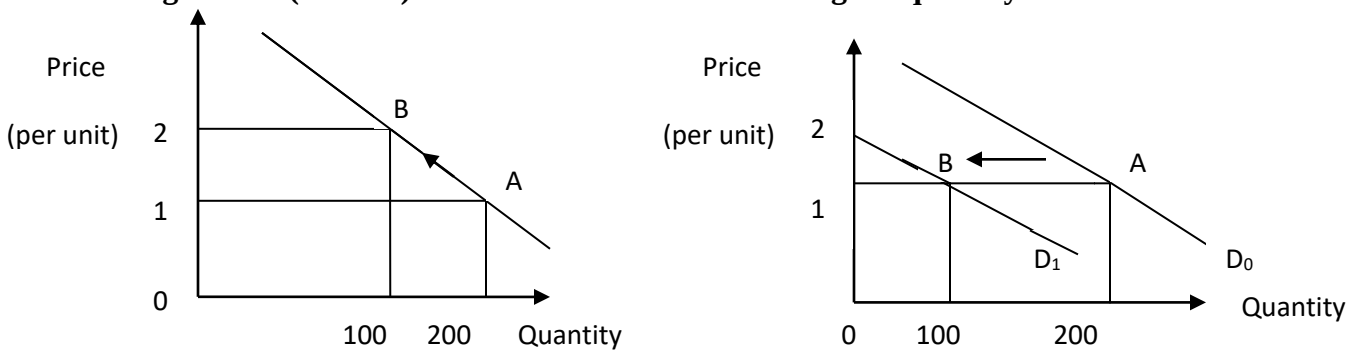
Price of other goods: Because people make their buying decisions based on the price of related goods, demand will be affected by the prices of other goods. Suppose the price of jeans rose from $25 to $35, but the price of khakis remained at $25. Next time you need shirts, you are apt to try khakis instead of jeans. They are substitutes. When the price of a substitute declines, demand for the other good whose price has remained the same will fall. Suppose the price of benzene falls. What will happen to the demand for a car? The lower cost of benzene increases the demand for a car. When the price of a good declines, the demand for its complement rises.

Tastes: "There is no accounting for taste". Of course, many advertisers believe otherwise. Changes in taste can affect the demand for a good with out a change in price.

Expectations: Finally, expectations will also affect demand. Expectations can cover a lot. If you expect your income to rise in the future, you are bound to start spending some of it today. If you expect the price of computers to fall soon, you may put off buying one until later.

Taxes and subsidies: Taxes levied on consumers increase the cost of goods to consumers and therefore reduce demand for those goods. Subsidies to consumers have the opposite effect.

Figure 2.2. (a and b) Shift in demand versus a change in quantity demanded



(a) Movement along a demand curve (b) Shift in demand

Individual and Market Demand Curves

Normally, economists talk about market demand curves rather than individual demand curves. A **market demand curve** is the horizontal sum of all individual demand curves. Market demand curves are what most firms are interested in. Firms do not care whether individual A or individual B buys their goods; they only care that someone buys their goods.

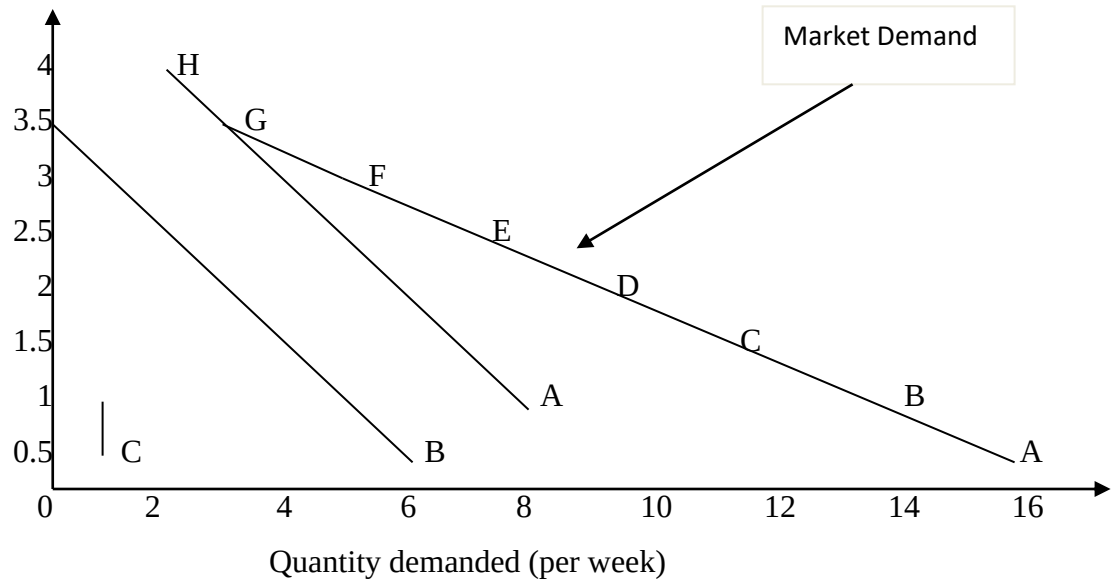
It is a good graphical exercise to add individual demand curves together to create a market demand curve (Figure 2.3). In this case it is assumed that the market consists of three buyers, A, B, and C, whose demand tables are given in Figure 2.3 (a). At a price of $3 each, A buys 4 goods; at a price of $2, A buys 6. C is an all or nothing individual. He/she buys 1 good as long as the price is equal to or below $1; otherwise he/she buys nothing. If you put plot C's demand curve, it is a vertical line. However, the law of demand still holds: As price increases, quantity demanded decreases.

The table (a) shows the demand schedules for A, B and C. Together they make up the market for the good. Their total quantity demanded (market demand) for goods at each price are given in column 5. As you can see in (b), A's, B's and C's demand curves can be added together to get the total market demand curve. For example, at a price of $2, C demands 0, B demands 3, and A demands 6, for a market demand of 9 (point D).

Table 2.1 demand schedules

	(1) Price (per unit)	(2) A's Demand	(3) B's Demand	(3) C's demand	(4) Market demand
A	0.5	9	6	1	16
B	1	8	5	1	14
C	1.5	7	4	0	11
D	2	6	3	0	9
E	2.5	5	2	0	7
F	3	4	1	0	5
G	3.5	3	0	0	3
H	4	2	0	0	2

(a) A demand table



(b) Adding demand curves

Figure 2.3 (a and b) From Individual Demand to a Market Demand

Figure 2.3 (b) shows three demand curves: one each for A, B, and C. The market, or total, demand curve is the horizontal sum of the individual demand curves. To see that this is the case, notice that if we take the quantity demanded at $1 by A (8), B (5), and C (1), they sum to 14, which is point B (14, 1) on the market demand curve. In practice, of course, firms don't measure individual demand curves, so they don't sum them up in this fashion. Instead, they estimate total demand. Still, summing up individual demand curves is a useful exercise because it shows you how the market demand curve is the sum (the horizontal sum, graphically speaking) of the individual demand curves, and it gives you a good sense of where market demand curves come from. It also shows you that, even if individuals don't respond to small changes in price, the market demand curve can still be smooth and downward sloping. This is because, for the market, the law of demand is based on two phenomena:

1. At lower prices, existing demanders buy more.
2. At lower prices, new demanders (some all or nothing demanders like C) enter the market.

2.1.3. Elasticity of demand and its determinants

The elasticity of demand measures the sensitivity of a change in a variable such as in this case, price, and the demand for that good. In other words, it measures the responsiveness of demand to the change in the variable.

A change in any of the demand determinants will cause a change in quantity purchased of a good per time period. A businessman needs a simple operational approach to find the contribution of the factors affecting demand for his product. The elasticity concept is one which meets this requirement of a businessman. The coefficient of elasticity of demand is the percentage change in the dependent variable (quantity demanded) divided by the percentage change in the

independent variable (demand determinant). The main measures are price elasticity of demand, income elasticity of demand and cross elasticity of demand.

Price elasticity of demand (e_d) = $\frac{Percentage change in quantity demanded}{percentage change in price}$

A general formula for calculating the coefficient of price elasticity, derived from the above definition is given by:

$$ed = -\frac{\frac{Q_2 - Q_1}{Q_1}100}{\frac{P_2 - P_1}{P_1}100} = -\frac{\frac{dQ}{Q_1}}{\frac{dP}{P_1}} = -\frac{dQ}{dP}\frac{P_1}{Q_1}$$

Therefore, in general, $e_d = -\frac{dQ}{dP}\frac{P}{Q}$ Where P is the original price, Q is the original quantity, dP is change in price, dQ is change in quantity. For a straight line demand curve $\frac{dQ}{dP}$ is the reciprocal of the slope of the demand line. The same formula can be used to find the elasticity of demand at a point on a non-linear demand curve, but in this case $\frac{dQ}{dP}$ is the reciprocal of the slope of a tangent to the curve at the point. The minus sign in the equation is conventional: it ensures that the elasticity of demand for normal goods is positive. Point elasticity formula measures the elasticity of demand at a point on the demand curve.

Elasticity can also be measured between two points on a demand curve. The measure of elasticity of demand between any two finite points on a demand curve is known as **arc elasticity**.

$$\text{Arc } e_d = -\frac{dQ}{dP}\frac{\frac{P_1 + P_2}{2}}{\frac{(Q_1 + Q_2)}{2}} = -\frac{dQ}{dP}\frac{P_1 + P_2}{Q_1 + Q_2}$$

Where P_1 and Q_1 represent the original price and quantity, and P_2 and Q_2 stand for the new price and quantity, respectively. The formula is used for both linear and non-linear demand curves. The price elasticity of demand can range from zero to infinity, but it can be divided into the following categories.

a) When the price elasticity of demand is greater than one, demand is elastic. For example, if a one percent change in price causes a three percent change in quantity of a good purchased, e_d =3 (3 percent divided by 1 percent), i.e.; demand is elastic.

b) When the price elasticity of demand is less than one, demand is inelastic. If a one percent change in price produces a one fifth percent change in quantity purchased, $e_d = \frac{1}{5}$. Demand is inelastic, due to the fact that $\frac{1}{5}$ is less than 1.

c) When the price elasticity of demand is equal to one, demand is unitary elastic. If a one percent change in price causes a one percent change in quantity purchased, e_d =1.

The price elasticity of demand has practical uses. A businessman trying to enhance his total sales revenue would like to know whether rising or reducing the price of the product will increase the

revenue. The concept of price elasticity of demand plays a crucial role in providing answer to this query. The relationship between price elasticity of demand and total revenue is described as follows:

Along a demand curve price and quantity demanded always move in opposite directions. How a change in price affects total revenue (price X quantity) depends upon elasticity of demand. If the seller reduces price and demand for his good is found to be elastic, total revenue will increase. This must be so because even though the price cut reduces total revenue, the rise in quantity demanded caused by the price cut, increases total revenue. i.e. demand is elastic means that when price is reduced by a given percent, quantity demanded increases by more than that percent, and as the consequence total revenue increases. If demand for a good is inelastic, then total revenue will fall, when its price falls. Hence, when demand is inelastic, it is better for the profit maximizing firm to increase price rather than reducing it. Finally, if demand for a good is unitary elastic, then a fall in price does not change total revenue. This means that when demand is unitary elastic if price falls by a given percent, quantity demanded increases by the same percent and as the result total revenue remains unchanged.

Table 2.2. Quantity demanded and sales revenue of good A

Price	7	6	5	4	3	2	1
Quantity demanded	1	2	3	4	5	6	7
Total sales revenue	7	12	15	16	15	12	7

Let us take numerical examples for the determination of the price elasticity of demand.

Example

From Table 2.2, calculate the price elasticity of demand when price equals 4.

Solution

Since we are asked to determine the price elasticity of demand at a point, on the demand curve, the appropriate formula is:

$e_d = - \frac{dQ}{dP}\frac{P}{Q}$

The slope of the line $\left(\frac{dP}{dQ}\right) = \frac{7-1}{1-7} = -\frac{6}{6} = -1$

The reciprocal of the slope $\left(\frac{dQ}{dP}\right) = -\frac{1}{1} = -1$

$e_d = - (-1)\frac{4}{4} = 1$

Example

Estimate the price elasticity of demand in the price range from 5 to 6.

Solution

Here, the relevant equation is the arc elasticity formula.

$e_d = - \frac{dQ}{dP}\frac{P_1 + P_2}{Q_1 + Q_2}$ We have already determined the slope of the line and its reciprocal is

$\left(\frac{dQ}{dP}\right)$, which equals minus one. Therefore, e_d = - (-1) $\frac{5+6}{3+2} = \frac{11}{5}$. The price and quantity demanded columns of Table 2.2. are reproduced in Fig 2.4.

From the figure it is clear that the price elasticity of demand at all prices above the price 4 is greater than 1 (demand is elastic). But at all prices below 4 demand is inelastic. At the point where price equals 4, demand is unitary elastic.

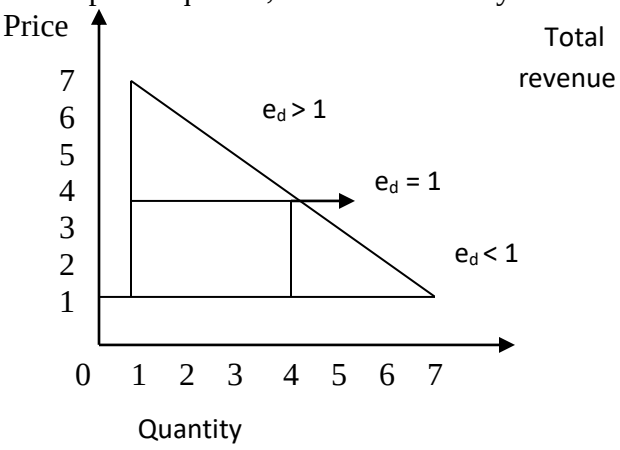


Figure 2.4 (a) Elasticity of demand on a linear demand curve

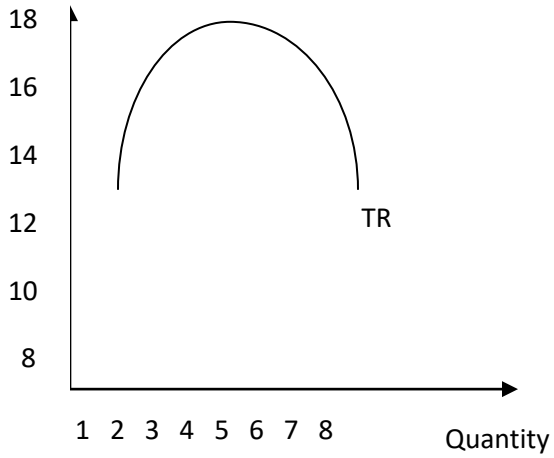


Figure 2.4(b) Total sales revenue curve

Table 2.2 and Figure 2.4 reveal the fact that if price decreases up to 4, total sales revenue increases because demand is elastic. At the price 4 total revenue is at its maximum and constant. Finally, if price decreases below 4 where demand is inelastic, total revenue decreases. Whether demand is elastic, inelastic or unitary elastic can also be determined by noting how a change in price affects revenue. This approach is known as total revenue test method. The method, however, does not give a value for the coefficient of elasticity of demand. According to this method,

a) If price and total revenue change in the same direction (example price decreases and total revenue decreases), demand is inelastic.

b) If price and total revenue change in opposite directions (example price decreases but total revenue increases), demand is elastic.

c) If total revenue does not change when price changes demand is unitary elastic.

There are three exceptional cases where the elasticity of demand remains the same (see fig 2.5). The three cases are known as perfectly inelastic, perfectly elastic, and unitary elasticity of demand.

a) Perfectly inelastic demand. Demand for a good is perfectly inelastic when there is no change in the quantity demanded whatever changes may occur in its price. It is represented by the vertical demand curve (**D_1)** in the figure; the price elasticity of demand for a perfectly inelastic demand curve equals zero.

b) Perfectly elastic demand. The horizontal demand curve (**D_2**) indicates perfectly elastic demand. A consumer will buy any desired quantities of the good at this price and nothing else at some other price. The price elasticity of demand equals infinite.

c) Unitary elasticity of demand. Demand for a good becomes unitary elastic when the percentage change in price and quantity demanded become the same. The rectangular hyperbola (**D_3)** in Figure 3.9 reflects the shape of the unitary elastic demand curve and its price elasticity of demand equals one.

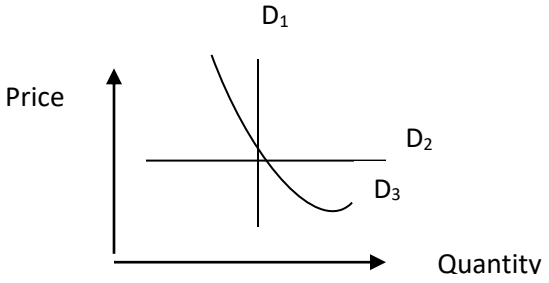


Figure 2.5 Three exceptional cases of elasticity of demand

Determinants of Price Elasticity of Demand

a) The availability of close substitutes. Demand for a good which has a close substitute (e.g. ; the brand kinds of soap Gulelle-555 and B-29, umbrella and rain-coat) tends to be elastic. Whenever there is a slight increase or decrease in the price of such goods it will be substituted by other close substitutes. But if the good has no close substitute, demand for it is inelastic.
b) Nature of the good. The consumption of necessities (e.g.; food-items, clothes) cannot be postponed, and hence demand for necessities is inelastic. On the other hand, consumption of luxury goods (e.g.; car, decoration-items) can be postponed, and demand for such goods tends to be elastic.
c) The proportion of income spent on the good. Demand for those goods on which a consumer spends very small proportion of his income will be inelastic (e.g.; salt, matches). For goods on which the proportion of income spent is high demand is elastic.
d) Time. When price of a good changes, a consumer does not respond immediately to the change in price. Over a period of time, however, the consumer is able to adjust his expenditure pattern to the price change. Therefore, demand is inelastic in the short-run but more elastic in the long-run.

Apart from the price of good, consumer's income is another basic determinant of demand for good. Income elasticity of demand (e_Y) shows the way in which a consumer's purchases of a good change as a result of change in his income, other things being constant.

$$e_Y = \frac{Percentage change in quantity demanded}{Percentage change in income}$$

Even though a change in income causes a change in quantity demanded, the two changes may not be of the same proportion. For example, if an individual has a 6% increase in his income, this does not mean he will buy 6% more of all the goods he used to buy. The consumer may buy more of some goods and less of some other goods because he can now afford something better and more expensive. Some goods may still be out of his reach, in which case he will buy nothing.

For all normal goods, income elasticity of demand is positive though the degree of elasticity varies in accordance with the nature of the good. Income elasticity of demand for inferior goods is negative. This is so because when the income of the consumer increases, he substitutes the superior good for the inferior one.

Demand for a good is also affected by the change in the prices of related goods. Cross-elasticity of demand (e_{AB}) measures the responsiveness of quantity demanded of good **A**, when there is a change in the price of some other good **B**.

$$e_{AB} = \frac{Percentage change in quantity demanded of\ A}{Percentage change in\ price of\ B}$$

The cross elasticity of demand is zero if the goods are unrelated; positive if the goods are substitutes; negative for complementary goods. The formulae for measuring the income elasticity and cross-elasticity of demand are the same as for measuring the price elasticity. The only difference is that the variables 'income' and 'price of related goods' are substituted for the variable 'price' in their respective formula.

2.1.4. Elasticity of demand of agricultural commodities

The demand for agricultural products is relatively inelastic. This is especially true for such basic crops as teff, wheat, corn, or beans. An important implication of this inelasticity is that even modest changes in supply, often brought about by weather patterns, can have large effects on the prices of those crops. This volatility in crop prices has been a feature of farming throughout all of history.

• Demand for most farm products is inelastic.
• People can consume only so much then they are satiated. Even if price drops they will not buy much more.

Consumers apparently put a low value on additional farm output compared with the value they put on additional units of alternative goods. Why is this so? Because one of the basic determinant of elasticity of demand is substitutability. When the price of one product falls, the consumer tends to substitute that product for other products whose prices have not fallen. Although people may eat more, they do not switch from three meals a day to, say, five or six meals a day in response to a decline in the relative prices of farm products. Real biological factors constrain an individual’s capacity to substitute food for other products.

The inelastic agricultural demand is also related to diminishing marginal utility. In a high income economy, the population is generally well-fed and well-clothed; it is relatively saturated with the food and fiber of agriculture. Consequently, additional farm products are subject to rapidly diminishing marginal utility. It takes very large price cuts to induce small increases in food and fiber consumption.

2.2. Theory of Supply

Supply and demand are the most-used words in economics. And for good reason. They provide a good off-the-cuff answer for any economic question. Try it.

Why are bacon and oranges so expensive this winter? Supply and demand.
Why are interest rates falling? Supply and demand.
Why can't I find decent wool socks anymore? Supply and demand.

Supply: In one sense, supply is the mirror image of demand. Individuals control the factors of production-inputs, or resources, necessary to produce goods. Individuals' supply of these factors to the market mirrors other individuals' demand for those factors. For example, say you decide you want to rest rather than weed your garden. You hire someone to the weeding; you demand labor. Someone else decides she would prefer more income instead of more rest; she supplies labor to you. You trade money for labor; she trades labor for money. Her supply is the mirror image of your demand.

For large number of goods and services, however, the supply process is more complicated than demand. For many goods there is an intermediate step in supply: individuals supply factors of production to firms.

Let us consider a simple example. Say you are a bread baker. You supply your labor to the factor market. The bakery demands your labor (hires you). The bakery combines your labor with other inputs like wheat powder, sugar, salt, oil, etc and produces bread (production), which it supplies to customers in the goods market. For produced goods, supply depends not only on individuals' decisions to supply factors of production but also on firms' ability to produce-to transform those factors of production into usable goods.

The supply of non produced goods is more direct. Individuals supply their labor in the form of services directly to the goods market. For example, an independent contractor may repair your washing machine. That contractor supplies his labor directly to you. Thus, the analysis of the supply of produced goods has two parts: an analysis of the supply of factors of production to households and to firms, and an analysis of the process by which firms transform those factors of production into usable goods and services.

2.2.1. The Law of Supply and Determinants of Supply

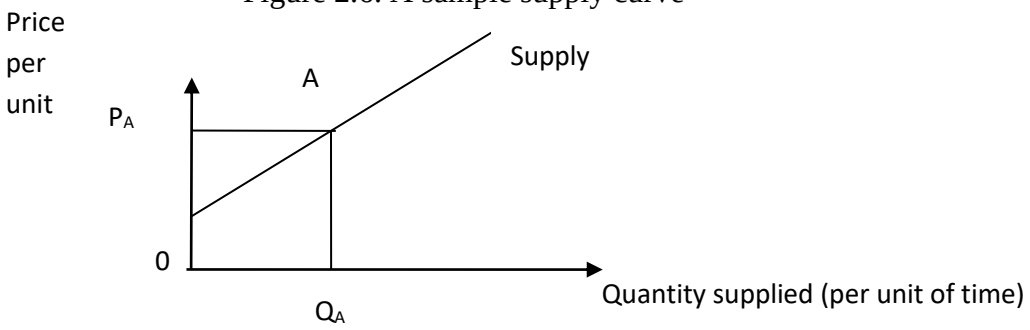
There is a law of supply that corresponds to the law of demand. The **law of supply** states:
Quantity supplied rises as price rise, other things constant. Or alternatively:
Quantity supplied falls as price falls, other things constant.
Price regulates quantity supplied just as it regulates quantity demanded. Like the law of demand, the law of supply is fundamental to the invisible hand's (the market's) ability to coordinate individuals' actions.

What accounts for the law of supply? When the price of a good rises, individuals and firms can rearrange their activities in order to supply more of that good to the market.

The supply curve

A supply curve is the graphical representation of the relationship between price and quantity supplied.

Figure 2.6. A sample supply curve



Notice how the supply curve slopes upward to the right. That upward slope captures the law of supply. It tells us that the quantity supplied varies directly-in the same direction-with the price.

As with the law of demand, the law of supply assumes other things are held constant. Thus, if the price of wheat rises and quantity supplied falls, you will look for something else that changed-for example, a drought might have caused a drop in supply. Your explanation would go as follows: Had there been no drought, the quantity supplied would have increased in response to the rise in price, but because there was a drought, the supply decreased, which causes prices to rise.

2.2.2. Shifts in Supply versus Movements along a Supply Curve

Supply refers to a schedule of quantities a seller is willing to sell per unit of time at various prices, other things constant.
Quantity supplied refers to a specific amount that will be supplied at a specific price.
In graphical terms, supply refers to the entire supply curve because a supply curve tells us how much will be offered for sale at various prices. "Quantity supplied" refers to a point on a supply curve, such as point **A** in figure 2.7.

The second distinction that is important to make is between the effects of a change in price and the effects of shift factors on how much of a good is supplied. Changes in price cause changes in quantity supplied; such changes are represented by a movement along a supply curve-the graphic representation of the effect of a change in price on the quantity supplied. If the amount supplied is affected by anything other than price, that is, by a shift factor of supply, there will be a shift in supply-the graphic representation of the effect of a change in a factor other than price on supply.

Shift Factors of Supply

Other factors besides price that affect how much will be supplied include the price of inputs used in production, technology, expectations, and taxes and subsidies. Let us see how:

Price of inputs: Firms produce to earn a profit. Since their profit is tied to costs, it is no surprise that costs will affect how much a firm is willing to supply. If costs rise, profit will decline, and a firm has less incentive to supply. Supply falls when the price of inputs rises. If costs rise substantially, a firm might even shut down.

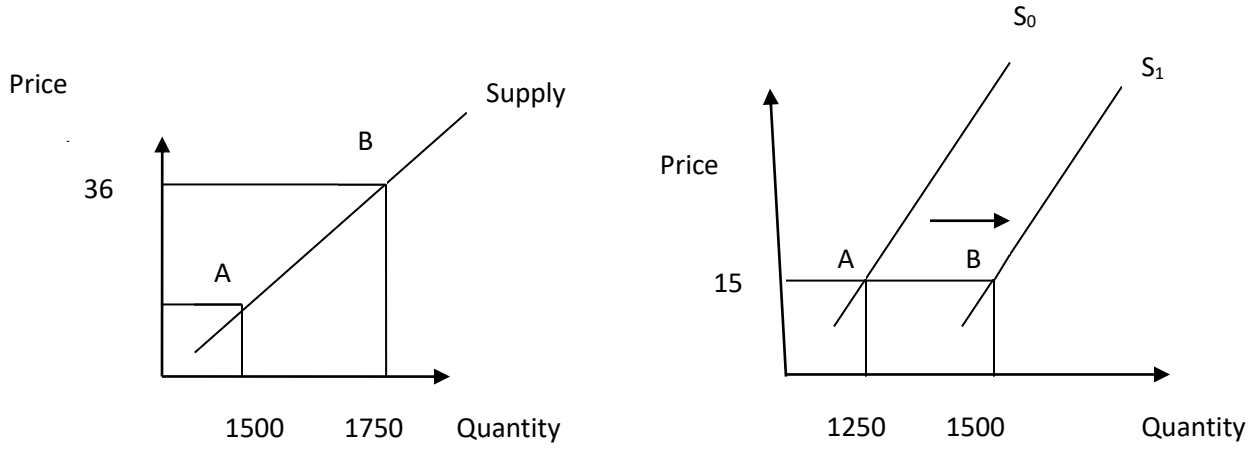
Technology: Advances in technology change the production process, reducing the number of inputs needed to produce a given supply of goods. Thus, a technological advance that reduces the number of workers will reduce costs of production. A reduction in the costs of production increases profits and leads suppliers to increase production. Advances in technology increase supply.

Expectations: Supplier expectations are an important factor in the production decision. If a supplier expects the price of her good to rise at some time in the future, she may store some of today's supply in order to sell it later and reap higher profit, decreasing supply now and increasing it later.

Taxes and subsidies: Taxes on suppliers increase the cost of production by requiring a firm to pay the government a portion of the income from products or services sold. Because taxes increase the cost of production, profit declines and suppliers will reduce supply. The opposite is true for subsidies-subsidies to suppliers are payments by the government to produce goods; thus, they reduce the cost of production. Subsidies increase supply. Taxes on suppliers reduce supply. These are not the only shift factors. As was the case with demand, a shift factor of supply is anything that affects supply, other than its price.

Figure 2.7. Shift in supply versus change in quantity supplied

A change in quantity supplied results from a change in price and is shown by a movement along a supply curve like the movement from point **A** to point **B** in (a). A shift in supply-a shift in the entire supply curve-brought about by a change in a non price factor is shown in (b).



(a) Movement along a supply curve (b) Shift in supply

The same "movement along" and "shift of" distinction that we developed for demand exists for supply. To make that distinction clear, let us consider an example: the supply of oil. In 1990 and 1991, world oil prices rose from $15 to $36 a barrel when oil production in the Persian Gulf was disrupted by the Iraqi invasion of Kuwait. U.S. oil producers, seeing that they could sell their oil

at a higher price, increased oil production. As the price of oil rose, domestic producers increased the quantity of oil supplied. The change in domestic quantity supplied in response to the rise in world oil prices is illustrated in Figure 3.5 (a) as a movement up along the U.S. supply curve from point **A** to point **B**. At $15 a barrel, producers supplied 1,500 million barrels of oil a day, and at $36 a barrel they supplied 1,750 million barrels per day.

Earlier, in the 1980s, technological advances in horizontal drilling more than doubled the amount of oil that could be extracted from some oil fields. Technological innovations such as this reduced the cost of supplying oil and shifted the supply of oil to the right, as shown in Figure 2.7 (b). Before the innovation, suppliers were willing to provide 1,250 million barrels of oil per day at $15 a barrel. After the innovation, suppliers were willing to supply 1,500 million barrels of oil per day at $15 a barrel.

Individual and market supply curves

The market supply curve is derived from individual supply curves in precisely the same way that the market demand curve was. To emphasize the symmetry, the three suppliers quite similar to the three demanders were taken. **A** (column 2) will supply 2 at $1; if price goes up to 2, she increases her supply to 4. **B** (column 3) begins supplying at $1, and at $3 supplies 5, the most he will supply regardless of how high price rises. **C** (column 4) has only two units to supply. At a price of $3.50 he will supply that quantity, but higher prices won't get him to supply any more.

The market supply curve's upward slope is determined by two different sources: by existing suppliers supplying more and by new suppliers entering the market. Sometimes existing suppliers may not be willing to increase their quantity supplied in response to an increase in prices, but a rise in price often brings brand-new suppliers in to the market.

2.2.3. Elasticity of supply and its Determinants

Just like the elasticity of demand, there is the concept of elasticity of supply. The elasticity of supply shows the responsiveness of quantity supplied when one of the supply determinants changes.

$$es = \frac{Percentage\,change\,in\,quantity\,supplied}{Percentage\,change\,in\,price}$$

i.e.; $es = \dfrac{\frac{dQ}{Q}}{\frac{dP}{P}} = \dfrac{dQ}{dP}\,\dfrac{P}{Q}$ Where $\frac{dQ}{Q}$ is percentage change in quantity supplied, and $\frac{dP}{P}$ is percentage change in price.

The price elasticity of supply may vary from zero to infinity. When the quantity of a good available for sale changes little with a big change in its price, supply is inelastic ($e_s < 1$**).** If the quantity of a good supplied changes by a greater amount than the change in price, supply is

elastic ($e_s > 1$**).** If the quantity of a good supplied changes by the same amount as the change in price, supply is unitary elastic ($e_s = 1$**).**
Similar to the elasticity of demand, there are three exceptional cases where the price elasticity of supply is always the same (see Figure 3.10). These cases are perfectly inelastic, perfectly elastic, and unitary elastic supply.
a) If e_s = **0**, supply is perfectly inelastic. It means there is no effect of price changes on quantity supplied. The supply curve will be vertical (S_1**).**
b) A unitary elastic supply curve ($e_s = 1$**)** passes through the origin as shown by S_2**.**
c) Lastly, if supply is perfectly elastic ($e_s = \infty$**)**, supply curve is horizontal (S_3**).** It means that infinite quantity is supplied at the ruling price.
b) A unitary elastic supply curve ($e_s = 1$**)** passes through the origin as shown by S_2**.**
c) Lastly, if supply is perfectly elastic ($e_s = \infty$**)**, supply curve is horizontal (S_3**).** It means that infinite quantity is supplied at the ruling price.

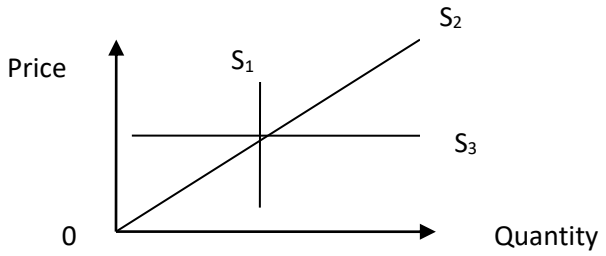


Figure 2.8. Three exceptional cases of elasticity of supply.

Determinants of the elasticity of supply

a) Length of time. The price elasticity of supply is influenced by the length of time which is required for adjusting supply to change in price. In the momentary time, while buyers and sellers are in the market, supply is perfectly inelastic. This means that quantity supplied is limited to the quantities already available in the market and it cannot be increased whatever price is offered. In the short-run, there is not enough time to vary all factors of production and only a limited change in quantity supplied may be possible by employing more variable inputs. Therefore, supply is more elastic in the short-run than in the momentary time. In the long-run, however, there is sufficient time to adjust all factors of production and here supply is much more elastic than in the short-run.

How long is the production process? Factories make cars very quickly nowadays with the production lines filled with robots. On the other hand, the length of time from the planting of the seed to the harvesting of the crop for most **agricultural products is very long**, relatively. The shorter the production period the more **elastic** supply will be as, again, the firm in question can respond more quickly to a change in the price level.

b) Availability of factors of production. If the factors of production are available in required amount, supply is relatively elastic. Otherwise supply is inelastic.

Does the industry have spare capacity? During a recession, there tends to be more **spare capacity** in firms and industries. This means that firms have the ability to increase output; perhaps some of their machines are currently unused. In these situations, firms will be able to respond quickly to a rise in the price of their product and so supply will be relatively **elastic**. If the economy, or the firm in question, is close to **full capacity**, then they will not be able to respond so quickly to a price rise so the supply curve will be relatively **inelastic**.

c) Possibility to produce alternative goods. If it is easy to switch over to the production of other goods supply will be elastic. If it is not easy to shift resources to the alternative production, supply is inelastic.

d) Nature of the good (Can the product be stored?). Some goods can be kept as stock very easily. You have probably seen the pictures of the huge areas of parked new cars on TV whenever there is a feature about the declining car industry on the news. If a firm has the ability to stockpile the product in question, then it can respond quickly to a rise **or** a fall in price by simply running down or piling up its stocks. In this case supply would be relatively **elastic**. Firms that sell perishables, for example (like bananas), cannot store old produce, and so their ability to respond to changes in the price level is reduced. Their supply will be relatively **inelastic** on this count. Many goods, especially **agricultural goods**, have an **inelastic supply** in the short run because producers are unable to increase supply significantly with short notice.

2.2.4. Market Equilibrium

When you have a market in which neither suppliers nor consumers collude and in which prices are free to adjust, the forces of supply and demand interact to arrive at equilibrium. In supply/demand analysis, equilibrium means that the upward pressure on price is exactly offset by the downward pressure on price. Equilibrium price is the price toward which the invisible hand drives the market. At the equilibrium price, quantity demanded equals quantity supplied. Equilibrium quantity is the amount bought and sold at the equilibrium price.
What happens if the market is not in equilibrium-if quantity supplied doesn't equal quantity demanded? You get either excess supply or excess demand, and a tendency for prices to change.
Excess supply: If there is excess supply (a surplus), quantity supplied is greater than quantity demanded, and some suppliers will not be able to sell their goods. All suppliers with excess goods will be thinking the same thing; the price in the market will fall. As that happens, consumers will increase their quantity demanded. So the movement toward equilibrium caused by excess supply is on both the supply and demand sides.
Excess demand: The reverse is also true. Say that instead of excess supply, there is excess demand (a shortage)-quantity demanded is greater than quantity supplied. There are more consumers who want the good than there are suppliers selling the good.

Price adjusts: When quantity demanded is greater than quantity supplied, prices tend to rise. When quantity supplied is greater than quantity demanded, prices tend to fall. Two other things to note about supply and demand are (1) the greater the difference between quantity supplied and quantity demanded, the more pressure there is for prices to rise or fall, and (2) when quantity demanded equals quantity supplied, the market is in equilibrium.

People's tendencies to change prices exist as long as there is some difference between quantity supplied and quantity demanded. But the change in price brings the law of supply and demand into play. As price falls, quantity supplied decreases as some suppliers leave the business (the law of supply). And as some people who originally were not really interested in buying the good think, "well, at this low price, might I do want to buy", quantity demanded increases (the law of demand). Similarly, when price rises, quantity supplied will increase (the law of supply) and quantity demanded will decrease (the law of demand).

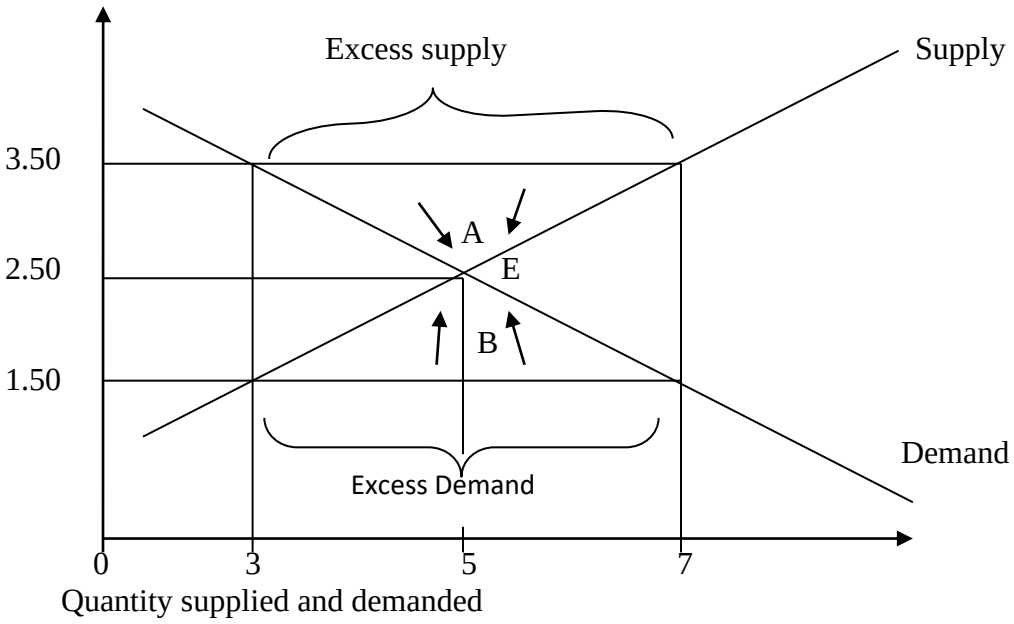
The graphical interaction of supply and demand

Figure 3.4 shows supply and demand curves for a good and demonstrates the force of the invisible hand. Let's consider what will happen to the prices of a good in three cases:
1. When the price is $3.50 each;
2. When the price is $1.50 each; and
3. When the price is $2.50 each.
1. When price is $3.50, quantity supplied is 7 and quantity demanded is only 3. Excess supply is 4. Individual consumers can get all they want, but most suppliers cannot sell all they wish; they will be stuck with a good that they would like to buy. Suppliers will tend to offer their goods at a lower price and demanders, who see plenty of suppliers out there, will bargain harder for an even lower price. Both these forces will push the price as indicated by **A** arrows in Figure3.6. Now let us start from the other side.
2. Say price is $1.50. The situation is now reversed. Quantity supplied is 3 and quantity demanded is 7. Excess demand is 4. Now it is consumers who can't get what they want and suppliers who are in the strong bargaining position. The pressures will be on price to raise in the direction of the **B** arrows in Figure 3.6.
3. At $2.50, price is at equilibrium: quantity supplied equals quantity demanded. Suppliers offer to sell 5 and consumers want to buy 5, so there is no pressure on price to rise or fall. Price will tend to remain where it is (point **E** in Figure 3.6). Notice that the equilibrium price is where the supply and demand curves intersect.

Figure 2.9. The interaction of supply and demand

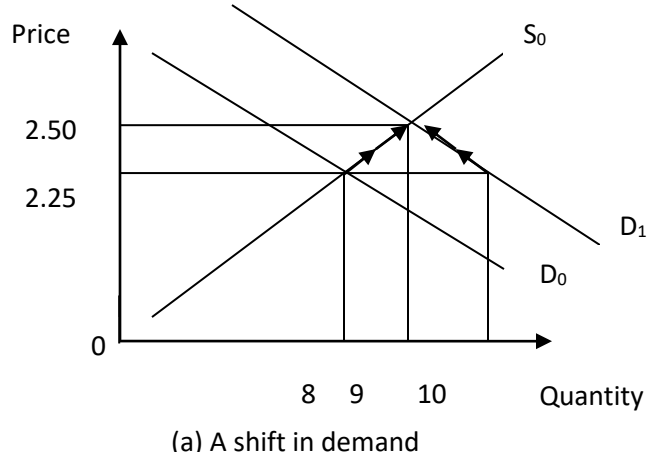
When there is excess demand there is upward pressure on price. When there is excess supply there is downward pressure on price. Understanding these pressures is essential to understanding how to apply economics to reality.
Table 2.3. The interaction of supply and demand

Price (per unit)	Quantity supplied	Quantity demanded	Surplus (+) shortage (-)
$3.50	7	3	+4
$2.50	5	5	0
$1.50	3	7	-4

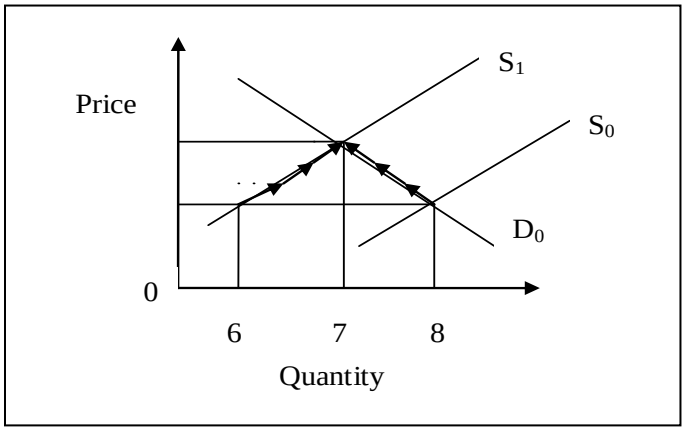


Shifts in Supply and Demand

Supply and demand are most useful when trying to figure out what will happen to equilibrium price and quantity if either supply or demand shifts. Figure 2.10(a) deals with an increase in demand. Figure 2.10(b) deals with a decrease in supply.



(a) A shift in demand



(b) A shift in supply

Let us consider again the supply and demand for a good. In figure 2.10(a), the supply is **S_0** and initial demand is **D_0.** They meet at an equilibrium price of $2.25 per unit and an equilibrium quantity of 8 units per week (point **A**). Now say that the demand for a good increases from **D_0** to **D_1.** At a price of $2.25, the quantity of a good supplied will be 8 and the quantity demanded will be 10; excess demand of 2 exists.

The excess demand pushes price upward in the direction of the small arrows, decreasing the quantity demanded and increasing the quantity supplied. As it does so, movement takes place along both the supply curve and the demand curve. The upward push on price decreases the gap between the quantity supplied and the quantity demanded. As the gap decreases, the upward

pressure decreases, but as long as that gap exists at all, price will be pushed upward until the new equilibrium price ($2.50) and new quantity (9) are reached (point **B**). At point **B**, quantity supplied equals quantity demanded. So the market is in equilibrium. Notice that the adjustment is twofold: the higher price brings about equilibrium by both increasing the quantity supplied (from 8 to 9) and decreasing the quantity demanded (from 10 to 9).

Figure 2.10(b) begins with the same situation that we started with in Figure 2.10(a); the initial equilibrium quantity and price are 8 units per week and $2.25 per unit (point **A**). In this example, however, instead of demand increasing, let us assume supply decreases-say because some suppliers change what they like to do, and decide they will no longer supply goods. That means that the entire supply curve shifts inward to the left (from $\mathbf{S_0}$ to $\mathbf{S_1}$). At the initial equilibrium price of $2.25, the quantity demanded is greater than the quantity supplied. Two more units are demanded than are supplied. (Excess demand=2).

This excess demand exerts upward pressure on price. Price is pushed in the direction of the small arrows. As the price rises, the upward pressure on price is reduced but will still exist until the new equilibrium price, $2.50, and new quantity, 7, are reached. At $2.50, the quantity supplied equals the quantity demanded. The adjustment has involved a movement along the demand curve and the new supply curve. As price rises, quantity supplied is adjusted upward and quantity demanded is adjusted downward until quantity supplied equals quantity demanded where the new supply curve intersects the demand curve at point **C**, equilibrium of 7 and $2.50.

Price controls

The government sometimes intervenes with the functioning of the price mechanism by establishing price ceilings and price floors. Price ceilings are maximum price fixation by law. That is, if the government imposes price ceilings, price is not legally allowed to increase above the fixed price. For example, if the equilibrium price determined by the two forces is Birr 2.00. But, if the government passes a law which fixes the price at Birr 1.50 this is known as a **price ceiling**. The economic consequence of a price ceiling is that of creating shortage. Shortage, in its turn, is likely to:

a) Ripe conditions for a black market and in the black market, usually, higher price is charged than the fixed price.

b) Call for non-price distribution like rationing which might be ensued by queuing, bribing and so on.

Price floors are minimum price fixation by law. In this case, price is not allowed to decrease below the fixed price. For example, if the government fixes price at Birr 2.50 (above the equilibrium price) such price control is known as **price floors**. The economic consequence of a price floor is surplus. Governments usually impose price floors on such goods as agricultural products (in the form of price supports) and wages (in the form of minimum wages).

2.3. The Utility Analysis

Why does a consumer buy a good or service? The answer is: because he obtains satisfaction from the possession of the good or service. In a technical term, a consumer buys a good or service because it has utility to him.

Utility is the amount of satisfaction to be obtained from a good or service at a particular time. For instance, the utility of orange is the satisfaction derived from consuming it at a particular time. There are some points to note: First, whether the good is useful or not, so far it gives satisfaction to the consumer, it has utility for him. Second, if a consumer wants something whether it is good or bad for him, it has utility for him. That is, it may satisfy socially immoral want, e.g. drug, alcoholism etc. Third, utility of a good varies from person to person and it varies for the same person from time to time. This implies that utility depends on the individual's own subjective estimate of the amount of satisfaction to be derived from a good or service.

2.3.1. Cardinal and Ordinal Utility

The early economists (like Jeremy Bentham-1800s) believed that utility is measurable either in terms of money or in simple numbers. According to them the unit of measurement of utility is 'util'. In their economic analysis they assumed that: a) one "util" equals one unit of money. b) Utility of money remains constant. This theory is known as the **cardinal** approach because it believes that utility can be assigned a cardinal number like 1, 2, 3,----. While classical and neo-classical economists maintain that utility is numerically measurable, other economists believe that utility being a subjective and abstract concept cannot be measured. This theory is known as **ordinal** approach. According to the ordinal approach it may not be possible for a consumer to express the utility of a good in absolute (cardinal) terms. But, it is always possible for a consumer to tell whether a good is more or less or equally useful to him compared to another. For instance, a consumer may not be able to say tea gives him 10 utilities and coffee 15 utilities. But he can always tell that coffee gives him more utility than tea.

Assuming that utility can be measured (cardinal approach), total utility is the sum of utilities obtained by a consumer from consuming the various units of goods and services. When a consumer consumes **n** different goods and services, his total utility (**U**) is the sum of the utilities derived from all of them.

$U = U_1+U_2+---+U_n = f_1(q_1) + f_2(q_2) + ---+f_n(q_n)$ where **q** is quantity.
Therefore, utility obtained from consuming a good or service depends exclusively on its quantities, other things remaining constant. The additional satisfaction received over a given period by consuming one more unit of good is known as **marginal utility**.

When is a consumer maximizing his utility (equilibrium is established)? A consumer with limited money income needs several goods and services for consumption at a time. How should he spend his limited income on various goods and services so that he gets maximum utility? A rational and utility maximizing consumer consumes goods in the order of their utilities. He picks up first the good which yields the highest utility followed by the good yielding the second highest utility, and so on. Therefore, consumer equilibrium is established 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 (used up completely). This is a utility-maximizing rule and can be expressed mathematically by,

$\frac{MU_A}{P_A} = \frac{MU_B}{P_B}$ Provided that $P_AQ_A+P_BQ_B+---+P_NQ_N = Y$

Where, MU_A - marginal utility of A,
MU_B - Marginal utility of B,
P_A - price of A,
P_B - price of B,
Q_A - quantity of A,
Q_B - quantity of B, and
Y - Consumer income

For any pairs of goods we can further write the equilibrium condition as follows:

$\frac{MU_A}{P_A} = \frac{MU_B}{P_B}$ Subject to the constraint that, $P_AQ_A+P_BQ_B = Y$

The principle of rational choice: The principle of rational choice is as follows: Spend your money on those goods that give you the most marginal utility (MU) per Dollar. The principle of rational choice is important enough for us to restate.

If $\frac{MU_X}{P_X} > \frac{MU_Y}{P_Y}$, Choose to consume an additional unit of good **X.**

If $\frac{MU_X}{P_X} < \frac{MU_Y}{P_Y}$, Choose to consume an additional unit of good **Y.**

By substituting the marginal utilities and prices of goods into these formulas, you can always decide which good it makes more sense to consume. Consume the one with the highest marginal utility per Dollar.

Table 2.4. Utility maximizing combination of two goods

Quantity purchased	Income (Birr) = 8.00			
	Good A: price (Birr) = 2.00		Good B: price (Birr) = 1.00	
	Marginal utility	MU_A/P_A	Marginal utility	MU_B/P_B
1	36	18	30	30
2	24	12	22	22
3	20	10	16	16
4	18	9	12	12
5	16	8	10	10
6	10	5	4	4

Table 2.4 reveals a utility maximizing rule. The table shows that an individual can either spend all the eight Birr income on good A or good B or divide these between the two. If the consumer decides to buy the first unit of good A, paying Birr 2, he will get 36 units of marginal utility. But, if instead of buying the first of A he spends the two Birr on good B, he will be able to purchase the first two units of B, receiving 52 (30+22) units of marginal utilities. Thus, he chooses to buy the first 2 units of B. Then, he should spend his next two Birr on the first of A which gives him

more marginal utilities (36) than that of the third and fourth of B (28 MU). The next two Birr should be spent on the third and fourth of B. If the last two Birr is spent on the fifth and sixth of B he obtains 14 MU. But the second unit of A would give him 24 MU. Therefore, as a rational consumer he prefers buying the second unit of A. The total utility obtained by the individual is 140 (the sum of the MU of the first two units of A and the first four units of B). This represents the maximum utility the consumer can obtain from all expenditures, and the two conditions for consumer equilibrium are simultaneously satisfied:

$\frac{MU_A}{P_A} = \frac{MU_B}{P_B}$ $\frac{24}{2} = \frac{12}{1}$ 12 = 12

$P_AQ_A+P_BQ_B = Y$ $2\times2+4\times1 = 8$ 8 = 8

That is, the MU received from the last money spent on A (12) equals the MU received from the last money spent on B (12), and the total expenditure of the consumer on A and B (Birr 8) equals his money income (Birr 8). Therefore, the utility maximizing combination of the two goods is 2 units of A and 4 units of B which gives him a total utility of 140 (36+24+30+22+16+12).

2.3.2. The law of diminishing marginal utility and the demand curve

How does marginal utility of a good change, if its quantity consumed increases? There is a basic law of economics known as the **law of diminishing marginal utility** which explains this. The law of diminishing marginal utility states that the marginal utility of a good declines as more of it is consumed, over any given period, other things equal.

Table 2.5. The law of diminishing marginal utility as applied to good A

Units of good A	Total utility	Marginal utility
First	8	8
Second	15	7
Third	21	6
Fourth	26	5
Fifth	30	4

Assuming that utility is measurable, the law of diminishing marginal utility is illustrated numerically in Table 2. We see that with the increase in the number of units consumed per unit time, the total utility increases at a decreasing rate. The diminishing marginal utility is shown in the second column. Each successive unit yields less and less extra utility to the consumer than the previous one. Total utility is obtained by adding the marginal utility figures as shown in the table. Marginal utility is declining continuously and becomes zero when total utility is at its maximum. After the level where total utility is at its maximum, marginal utility is negative and total utility falls. A rational consumer would never go beyond the zero level of marginal utility while consuming a good.

The law of diminishing marginal utility is shown graphically in Figure 1. Part **a** of the figure shows the total utility curve which is rising but at a decreasing rate up to point **P** (the highest point of total utility). At point **P** it is constant and starts to fall thereafter. Part **b** of Figure1 show

the corresponding marginal utility curve, which is downward sloping. This proves the law of diminishing marginal utility. Geometrically, marginal utility is the slope of the total utility curve. When total utility is at its highest point (**P**), marginal utility equals zero. For the declining portion of total utility curve marginal utility becomes negative.

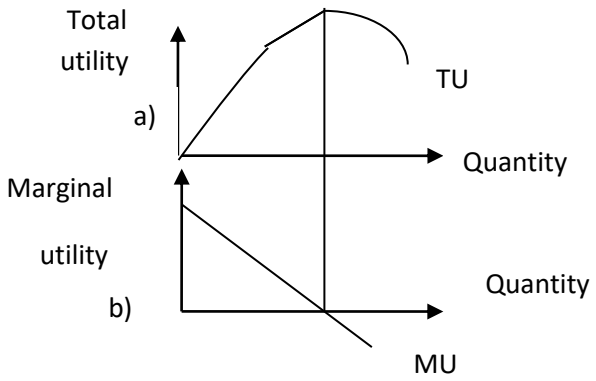


Figure 2.11. The relationship between total utility and marginal utility

2.3.3. Indifference Curve Analysis

The indifference curve analysis is an alternative approach of consumer behavior analysis. Here the emphasis is on comparing different utility levels instead of measuring them through cardinal scale. An indifference curve is a locus of points each showing a different combination of two goods which yield the same satisfaction to the consumer. Example, Table 2.6. Shows combinations of two goods, between which a consumer is indifferent. If the combinations **h, i, j, k, l,** in the table are plotted and joined by a smooth curve, the resulting curve is known as **indifference curve** (Figure 2.12).

Table 2.6. Indifference schedule of good A and B

Combination	Units of good A	Units of good B
H	24	4
I	12	8
J	8	12
K	6	16
L	5	20

How much of one good could the consumer give up in return for an additional unit of the other good, while enjoying the same level of utility? The answer to this query is as follows. At point **h**, in the table and the figure, the consumer has a large amount of good **A** (24 units) and small amount of **B** (4 units). This implies that the marginal utility of **B** is high, and the marginal utility of **A** is low to the consumer. It follows that at the margin (in terms of the marginal utility), **B** is very valuable and **A** is less valuable to him. Therefore, the consumer is willing to sacrifice 12 units of **A** to receive 4 units of **B**, in going from point **h** to **i.** As he moves from combination **h** to **i** the quantity of **A** decreases and that of **B** increases. As the amount of **A** decreases its marginal utility increases and as the amount of **B** increases, the marginal utility of additional units of **B** decreases (the law of diminishing marginal utility). Thus, in moving from point **i** to **j** the consumer is willing to give up only 4 units of **A** to obtain 4 units of **B**, and so on. This means that in Figure_ moving down the curve the consumer is substituting one good for the other but

maintaining a constant level of utility derived from them. The rate at the margin, at which the consumer substitutes one good for the other, so as to remain equally satisfied, is the slope of the indifference curve. It is known as the marginal rate of substitution.

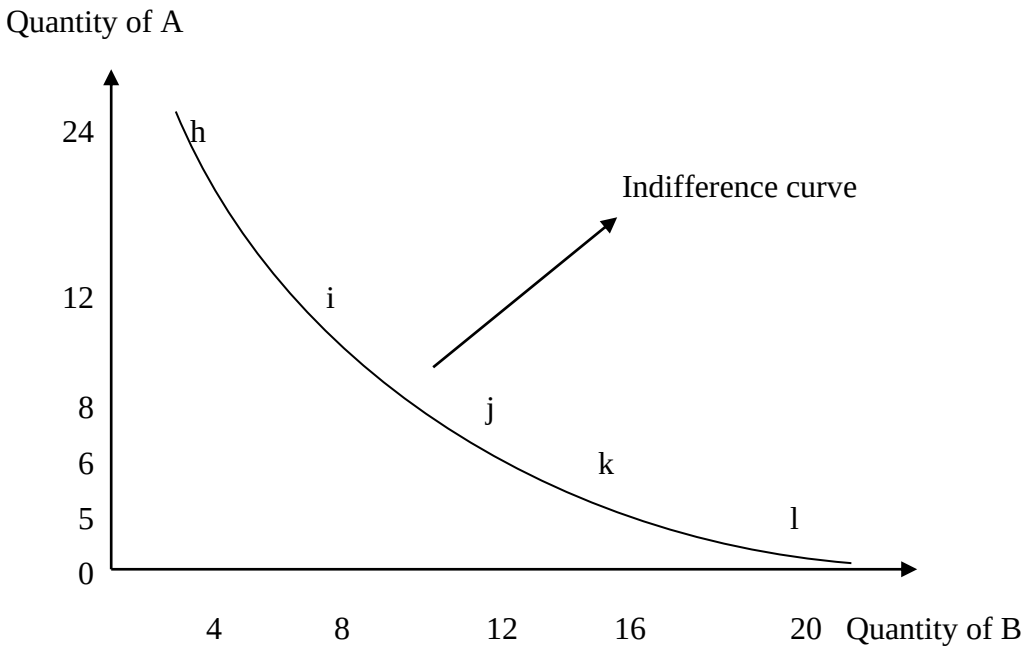


Figure 2.12.A consumer's indifference curve

Since the consumer is ready to give up less and less quantities of **A** to get additional units of **B**, the marginal rate of substitution (**MRS**) diminishes when he moves down the indifference curve. That is **MRS** decreases because in most cases no two goods are perfect substitutes for one another. If the two goods are perfect substitutes, the indifference curve will be a straight line with a negative slope and constant **MRS**. In symbolic form the **MRS** is denoted by,

$MRS_{BA} = -\frac{dA}{dB}$ Where MRS_{BA} is marginal rate of substitution of **B** for **A**, dA is change in quantity of good **A**, dB is the change in quantity of good **B**,

For example, in Figure_ on the indifference curve,

a) In moving from **h** to **i**, $MRS_{BA} = -\frac{12}{4} = -3$

The consumer is willing to give up 3 units of **A** in order to obtain one more unit of **B**.

b) In moving from **i** to **j**, $MRS_{BA} = -\frac{4}{4} = -1$

The consumer is willing to give up 1 units of **A** in order to obtain 1 unit of **B**.

If the consumer reduces the consumption of **A**, while the consumption of **B** kept constant, he will move from **i,** to **m** in Figure_. But **m** lies on a lower indifference curve. Hence, he must consume more of good **B** by moving from **m** to **j** to come back to the initial indifference curve, so as to leave his satisfaction unchanged. When the consumer goes from point **i** to **m**, the total loss in satisfaction equals the fall in quantity of **A** (dA) multiplied by the marginal utility of **A** (MU_A). Similarly, when he moves from point **m** to **j**, his total gain in satisfaction equals the increase in quantity of **B** (dB) multiplied by the marginal utility of **B** (MU_B). In order for the consumer to

remain on the same indifference curve, the satisfaction he looses by moving from **i** to **m** must be equal to the satisfaction he gains by moving from **m** to **j**, i.e;

$$-dA \times MU_A = dB \times MU_B \qquad -\frac{dA}{dB} = \frac{MU_B}{MU_A} = MRS_{BA}$$

The indifference curve in Figure_ relates to a fixed level of satisfaction. However, the combinations along this curve are not the only combinations of the two goods. In order to show consumer tastes at higher and lower levels of satisfaction (other combinations of the two goods), we have to draw other indifference curves. In the indifference plane, it is always possible to locate any two or more points indicating the same satisfaction. It is, therefore, possible to draw a number of indifference curves shown in Figure_. The set of indifference curves **I_1**, **I_2**, **I_3**, and **I_4** is known as indifference map. On a particular indifference curve the level of satisfaction remains the same. But the level of satisfaction differs on different indifference curves. The farther the indifference curve from the origin, the higher the utility it represents.

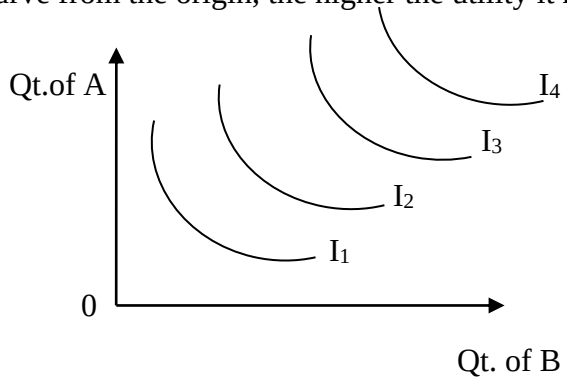


Figure 2.13.An indifference curve Map

Basic characteristics of indifference curve

a) An indifference curve slopes downward from left to right.
b) An indifference curve is convex to the origin.
c) Indifference curves do not intersect nor are they tangent to one another.

2.3.4. The Budget Constraint

Indifference curves show consumer's desire to buy different combinations of two goods. A consumer's desire is not enough to determine the equilibrium position. In other words, Indifference curves by themselves cannot tell us which combination is to be chosen. In addition to his, the consumer must have the capacity to pay for the goods. The consumer's capacity which is indicated by the budget line is determined by his income and the prices of the two goods. A utility maximizing consumer would like to reach the highest possible indifference curve on his indifference map. On the other hand, the income of the consumer which is assumed to be fixed sets limit to his choice (does not allow him to buy any desired quantities). This is known as budget constraint expressed in the following budget equation:

$$P_AQ_A+P_BQ_B = Y \qquad Q_A = \frac{Y}{P_A} - \frac{P_B}{P_A}Q_B, \quad Q_B = \frac{Y}{P_B} - \frac{P_A}{P_B}Q_A$$

So, from the above equation, if the values of $Y, P_A, P_B,$ are known, Q_A and Q_B can be easily calculated. For example, if the consumer buys only good A (Q_B =0), then $Q_A = \frac{Y}{P_A}$, and if he buts only good B (Q_A =0), then $Q_B = \frac{Y}{P_B}$. These are the two extremes of the consumer's income and expenditure relationship. In fact, there are other possible combinations of **A** and **B** between the two extremes. If the values of Q_A and Q_B obtained in the above mentioned procedures are plotted on **X** and **Y**-axis, they produce a downward sloping line known as **budget line** (see Figure 1 &2). A budget line is a boundary of consumption for the two goods. It divides the area between **X** and **Y**-axis into two parts, i.e.; feasibility area (the area outside the budget line).

For example, point **W** is in the non- feasibility area which means that given the money income of the consumer and the fixed prices of **A** and **B** it is not possible to buy the corresponding quantities of the goods at that point. On the other hand, a point like **V** is in the feasibility area which means that a consumer does not spend all his income on the two goods.
That is, buying the corresponding quantities of **A** and **B**, he is left with some unspent income.

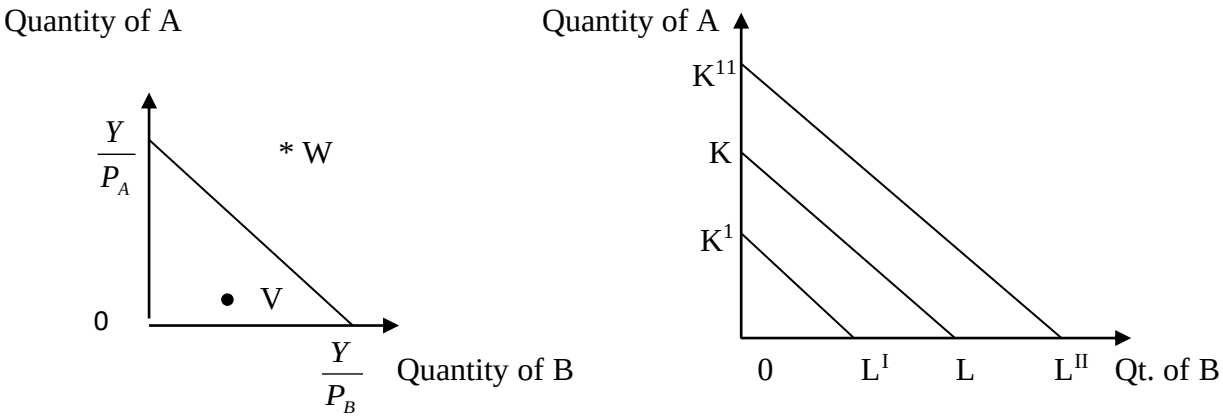


Figure 2.14. Consumer's budget line **Figure 2.15. Effects of changes in the consumer's income**

Changes in consumer income or prices of the goods cause shift in position of the budget line. Suppose the original budget line is **KL** in Figure 2.14. If the consumer's money income decreases (price kept constant), there will be a parallel leftward shift of the budget line, with no change in its slope. That is, after the fall in income the budget line is **K'L'**. On the other hand, if income increases, the budget line shifts to **K"L"**. The slope of the budget line KL in Figure 2 equals:

$\frac{dQ_A}{dQ_B} = \frac{0K}{0L}$, But $0K = \frac{Y}{P_A}$, and $0L = \frac{Y}{P_B}$ (see Figure2.14)

Therefore, $\frac{0K}{0L} = \frac{\frac{Y}{P_A}}{\frac{Y}{P_B}} = \frac{Y}{P_A} \times \frac{P_B}{Y} = \frac{P_B}{P_A}$

Hence, the slope of the budget line is the same as the ratio of the two prices. We may now consider the effect of changes in the prices of the two goods with a constant income. Changing both the prices simultaneously is a complex situation for analysis. Therefore, we just keep the price of **A** constant and illustrate the effect of the change in the price of **B.** If the price of **B** falls, the budget line **KL** shifts to **KM**, in Figure 3. This means that when $\mathbf{P}_B$ falls, the same income enables him to buy more of **B**. That is, quantity of **B** increases from **0L** to **0M** along the horizontal axis.

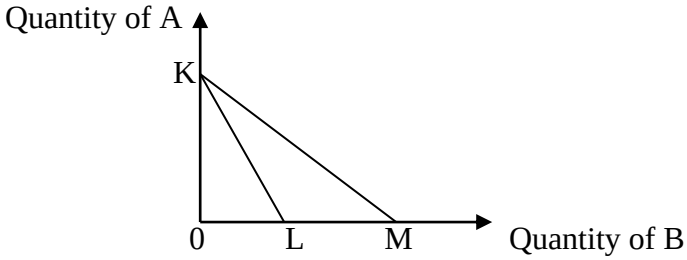


Figure 2.16. The effects of a fall in price on the budget line

2.3.5. Consumer Equilibrium (Combining indifference curves and budget constraint)

How is the equilibrium position determined in the indifference curve approach? A consumer whose objective is utility maximization would always attempt to reach the highest attainable indifference curve permitted by the budget line. We can now determine the consumer's equilibrium position by combining the indifference map and the budget line as shown in Figure 2.17.

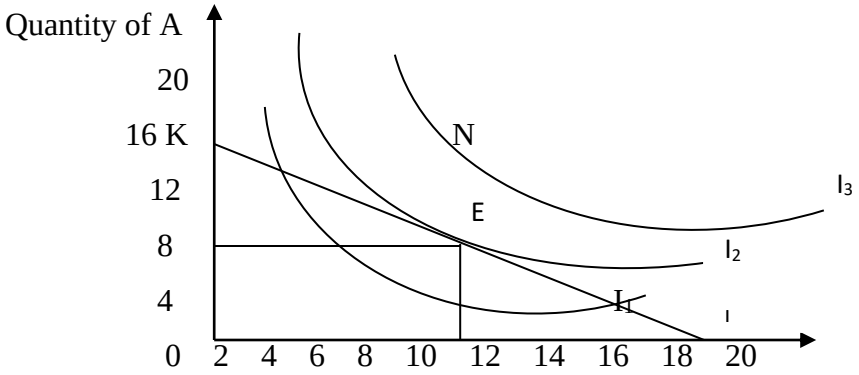


Figure 2.17. Consumer equilibrium

Which of the attainable combinations of good **A** and good **B** will the consumer most prefer? As a rational consumer, he would prefer that combination which yields the highest satisfaction. The utility maximizing combination will be the one lying on the highest attainable indifference curve (where the consumer desire coincides with his capacity). In Figure 2.17 the consumer's equilibrium combination of **A** and **B** is at point **E**. Here, the budget line is tangent to the

indifference curve **I_2**. At the point of tangency (**E**), the slope of the indifference curve = the slope of the budget line.

Or $\frac{MU_B}{MU_A} = \frac{P_B}{P_A}$ ----At equilibrium point **E**, the consumer consumes **Q_A** of good **A** and **Q_B** of good **B**, which yield him maximum satisfaction.

In the analysis of Figure 2.15 we said that a change in consumer income, the prices of the two goods constant, shifts the budget line either to the right or to the left. The shift in the position of the budget line, in its turn, affects the equilibrium position. For example, in Figure_ at equilibrium point **E** the budget line **KL** is tangent to **I_1**. But following the rightward shift of the budget line, to **K'L'**, equilibrium position changes from **E** to **E'** (on the higher indifference curve).Other things equal, the line joining the successive equilibrium points at various level of money income is known as the **income consumption curve**.

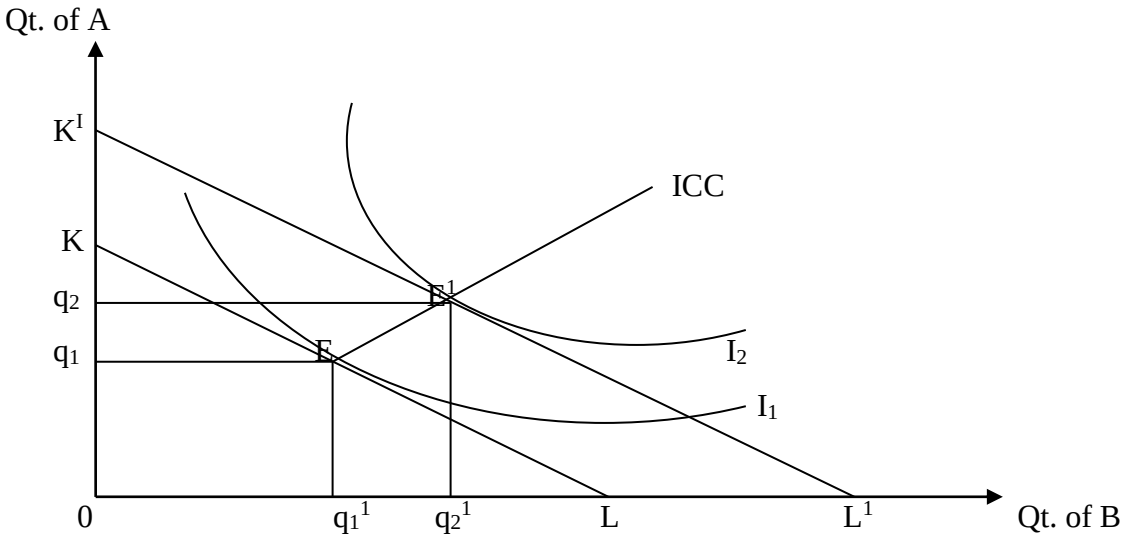


Figure 2.18. Derivation of income consumption curve

A rise in income may or may not increase demand curve for a good. For normal goods, as income increases, demand increases. Therefore, the income consumption curve is positively sloped for normal goods. For other goods, the income consumption curve bends backward over certain range of income.

CHAPTER THREE
Theory of Production and Cost in Agricultural Firms

3.1. Theory of Production: Short-run and Long-run

3.1.1. Production: is the process of changing economic resources in to output.
It is transformation of input into output ex, change in form (raw material changed in to finished goods), change in time (storage) and change in place (transportation).
Input **production** output
So production is the process of output and its distribution as well as storage of tangible goods. But not only tangible goods it also considers intangibles.
Production activities can be generally classified in to three:
A). Primary production: this involves the first direct process in which goods originate. Such as, agriculture, Fishing and mining are typical examples.
B). Secondary production: this type of economic activity refers to the process of raw materials in to final outputs.
C). Tertiary production: the aerial interchange of items which carried out in the form of trade and tourism as well as all forms of provision of services is a tertiary production. The provision of services may be direct or indirect. Direct services are of personal in nature such as those services provided by teachers, engineers, lawyers, doctor's technicians. Whereas indirect services are of commercial in nature which are more or less concerned with the distribution of goods from person to person, or from business to business, or from one area to another.

3.1.2. The role of the firm

A key concept in production is the firm. A **firm** is an economic institution that transforms factors of production into goods and services. Firms:
a) Organize factors of production,
b) Produce goods and services, and/or
c) Sell produced goods and services to individuals, businesses or government.
Which combination of activities a firm will undertake depends on the cost of undertaking each activity relative to the cost of subcontracting the work out to another firm. When the firm only organizes production, it is called a virtual firm. Virtual firms organize the factors of production and subcontract out all production. While most firms are not totally virtual, more and more of the organizational structure of businesses is being separated from the production process. As cost structures change because of technological advances such as the internet, an increasing number of well-known firms will likely concentrate on organizational instead of production activities.

3.1.3. Time period of production: the fixity and variability of an input depends on the length of time period under consideration.
Economists classify time period in two categories:
a). short run: is that period of time in which some of the firms inputs are fixed.
b). long run: is that period of time in which all inputs can be changed. No fixed inputs in the long run.

3.1.4. Production function

It is the physical relationship between input and output. The relationship shows the rate of transformation of input to output.

The production function is generally written in the form of equation:

$Y = f(X_1, X_2, X_3, \text{----------}, X_n)$

Where; Y= output level, Xi= inputs in the production process; f= function

Production function rest on two main assumptions:

1. Technology is invariant (fixed): if the technology changes it would result in an alteration of the input- output relationship depicted by the production function.
2. It s assumed that firms utilize their inputs at maximum level of efficiency.

There are numerous relationships between the resources and farm products, both simple and complex. The major production relationships fall under three categories;

1. Factor–product relationships or input-output relationship-production function: this relationship is concerned with resource allocation to optimum production. The choice indicator to be used is price ratio.

2. Factor-factor or input-input relationship or input combination: this relationship is concerned with minimizing cost at a given level of output. The choice indicators are price ratio and substitution ratio.

3. product-product relationship or output-output relationship or enterprise combinations: this relationship is concerned with optimum combination of outputs for a given input level. The choice indicators are price ratio and substitution ratio.

Production function can be expressed in terms of:

- Tabular form
- A graph
- An equation/ algebraic to specify maximum output rate from a given amount of inputs used.

A) Table/tabular presentation or tabulation (Table 3.1)

Units of input (labor in hour)	Capital	TPP	APP	MPP
1	5	5	5	5
2	5	11	5.5	6
3	5	18	6	7
4	5	25	6.25	7
5	5	30	6	5
6	5	32	5.3	2
7	5	32	4.3	0
8	5	28	4.5	-4

MPP (Marginal physical products): Marginal means additional or extra .MPP is the additional or extra TPP-produced by using an extra unit of input. It requires measuring changes in both out-put and input.

Marginal physical product is calculated as: $MPP = \frac{\Delta TPP}{\Delta X}$

The numerator is the change in TPP caused by a change in the variable input. The denominator is the actual amount of change in the input. Marginal physical product can be positive or negative.

It can be also zero if change in input level causes no change in TPP.A negative MPP indicates too much variable input is being used relative to the fixed input/s and this combination depresses TPP.

B) Mathematical presentation of production function

y = f(x) where f = function, y = dependent variable (output resulting from the production process); x = an aggregate independent variable (input that we change); $y = f(x_1|x_2$ ---------- $x_n)$ where x_1 = the variable input, (Eg. fertilizer) and x_1,x_2 = all fixed inputs for short period of time (e.g. land, labor, capital, etc.)

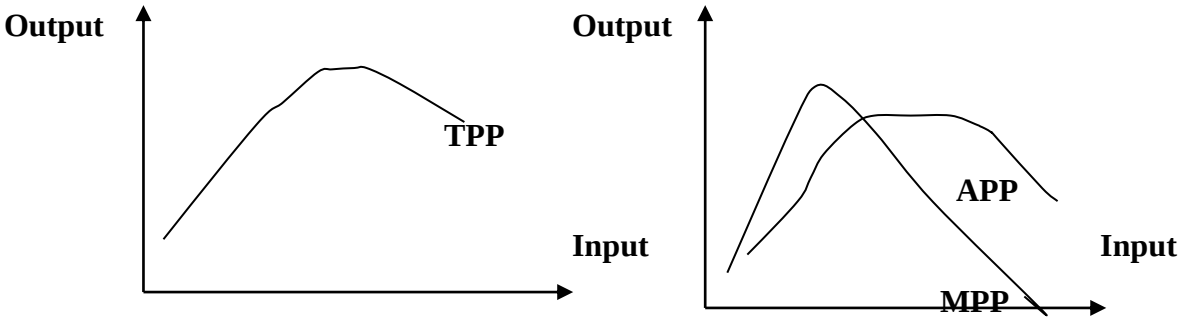
Therefore, the variation of y depends on the variable input x_1.

Example by considering a linear production function: y = a+bx or y = 20+4x

Where, a indicates the amount of y to be produced with non use of x

b indicates the slope of linear curve or marginal product we get from use of one additional unit of input x.

C) Graphical Illustration of a production function



Economic theory gives us two types of factor-product relationship in production function.

1. Proportionate relationship: In this relationship - the physical input-output relationship with one variable input with fixed inputs. i.e., some inputs are fixed, while quantities of other inputs vary showing short-run relationship.
2. Scale relationship: In this relationship, all inputs vary no fixed input. It is for a longer period, showing long run relationship.

3.1.5. Stages of production

Under proportionate relationship, there are three stages of production

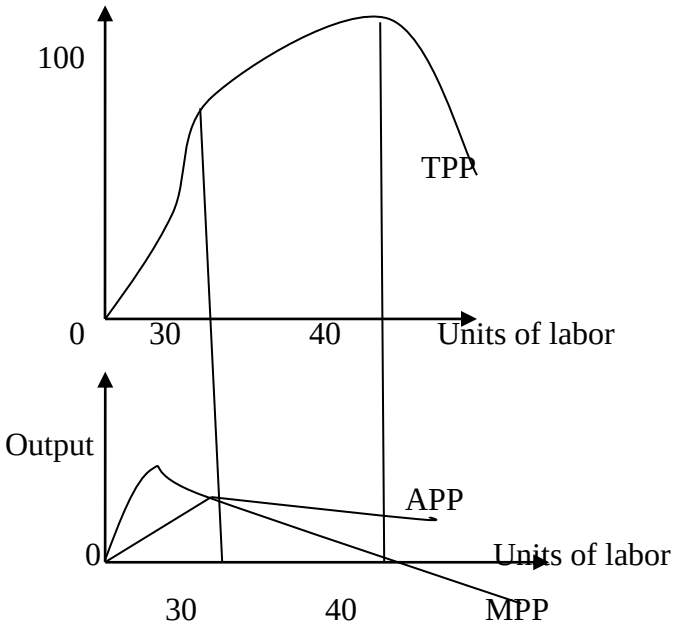


Figure 3.1. Three stages of production

Stage -I

- it extends from he origin to the point of APP at maximum(MPP=APP)
- more of fixed input, less of variable inputs comparatively(fixed inputs not full utilized)
- Ep>1 to Ep=1
- Uneconomical region

Stage-II

- starts from MPP=APP(APP maximum) and ends up where MPP=0(TPP at maximum)
- 0=Ep<1
- Economical region that Economic optimum is found
- We can't pin point the optimum production point by looking at the physical output only.

Stage -III

- starts ,where MPP=0 or TPP at its maximum
- Ep<0
- more of variable inputs
- Uneconomical region.

Relationship between TPP, APP and MPP

TPP Vs MPP	MPP Vs APP
♦ When MPP increases TPP increases at increasing rate ♦ When MPP decreases but greater than zero, TPP increases at decreasing rate ♦ When MPP=zero, TPP is at its maximum ♦ When MPP< zero, TPP declines	♦ When MPP greater than APP, APP is increasing ♦ When MPP is equal to APP, APP is at maximum ♦ When MPP is Less than APP, APP is decreasing

3.1.6. The Law of diminishing returns: The law stated that, if increasing amounts of one input are added to a production process, while all others inputs are held constant the amount of output added per unit of the variable input will eventually decrease. The law is a generalization based on experience that the use of increased inputs leads to less than proportionate increase in output.
The term diminishing return can be described in terms of either physical production or economic returns. TPP, APP or MPP values begin to diminish or decline after some point as more of the variable input is used. The law takes in to consideration total, marginal and average product curves for a variable input (x) in combination with the fixed inputs, the relationship is expressed through this common "**law of diminishing marginal return**". For diminishing marginal returns **to exist**, **one or more fixed inputs** must be used in the production process in addition to the variable input.

How Much output to produce under the law of diminishing return?
(Rules of profit Maximization)

To answer this question directly requires the introduction of two new marginal concepts namely Marginal Revenue (MR) and Marginal Cost (MC**).**

Marginal revenue (MR): is defined as the change in income or the additional income received from selling one more unit of output. It is calculated from the equation

$$MR = \frac{\Delta \text{ total revenue}}{\Delta \text{total physical product}}$$

Total Revenue (TR) is the same as total income. Total Revenue (TR) is used in place of total value product when discussing output levels**.**

MR: is constant and equal to the price of the price of output. The additional income received from selling one more unit of output will equal the price received for that output.

Marginal Cost (MC): is defined as the change in cost or the additional cost incurred from producing another unit of output. It is calculated from the equation:

$$MC = \frac{\Delta \text{ total input cost}}{\Delta \text{ Total physical product}}$$

Firms maximize profit: The firm plays the same role in the theory of supply that the individual does in the theory of demand. The difference is that whereas individuals maximize utility, firms maximize profit.

Profit = Total revenue –Total cost

The decision rule-MR and MC are compared to find the profit maximizing output level. When MR is greater than MC, the additional unit of output increases profit as the additional income exceeds the additional cost. Conversely, if MR is less than MC, producing the additional unit of output will decrease profit. At the output level where MR=MC, profit will be at its maximum

level There is only one profit maximizing combination of input and output for a given production function and a given set of prices. Once either the optimum input or output level is found the other value can be determined from the production function.

3.1.7. Elasticity of production (Ep): it is a measure of the responsiveness of production with a minor change in input. It is expressed as a percentage change in output as compared to a percentage change in input.
It is computed as:

$$E_P = \frac{\%\ \Delta\ in\ output\,(Y)}{\%\ \Delta\ in\ input\,(x)}$$

$$E_P = \frac{\frac{\Delta Y}{Y}}{\frac{\Delta X}{X}} = \frac{\Delta Y}{\Delta X} * \frac{X}{Y} = \frac{Mpp_x}{App_x}$$

3.2. Theory of Costs: Short run and Long run

3.2.1. Basic Concepts: cost, expense, classification of costs, and cost curves

Cost generally refers to expenses on goods and productive services to be produced.
They are expenses incurred in organizing and carrying out the production process in business.
Costs and expenses are different things-the later are

Expenses-are money incurred in the production process (Accounting)
Costs-represent efforts, sacrifice, and exertion of human being (Economic)
So, managers should be familiar with all of the costs and their classifications and understand which costs are important in decision making at a particular time.
Production cost plays a great role in decision making process because profit depends on amount of cost.
Classification of costs
A) Implicit and explicit
Explicit /cash/accounting cost-these are the actual monetary payments or cash outlays that business firms make to outsiders.
They are expenses which are actually paid by the firm. These costs appear in the actual recording of the firm.
E.g. Interest payment on borrowed fund, salary payments for workers, payments made for raw materials.
Implicit /non- cash/economic costs-these are the value of non-purchased inputs owned and used by a firm in his own production activities.
They are costs of the firms self-owned and self –employed resources in carrying out productive activities.
E.g. the salary of the owner-manager, estimate rent of building, family labor
The value of these self-owned inputs should be estimated from what they could earn in their best alternative use.

B) Material/service cost & operational cost

Material costs-are costs of materials which are totally or partially involved in the production process.

E.g. Crop production - land, land improvement, building, human labor, animal power, machinery, seed, fertilizer.

Livestock production –cost of building, fencing, electricity, water, feeding, machinery & tools, maintenance & depreciation.

Operational costs-costs incurred in the process of production

E.g. **For crop production**-cost of plowing, weeding, fertilization, irrigation, harvesting, etc

For livestock production-cost of feeding, grazing, breeding practices

C) Economic cost and Accounting cost

- **Economic cost/opportunity cost**-the opportunity cost of a good or service is the value of the best alternative that is foregone in order to produce the good undertake the service activity. Economic cost includes both explicit and implicit costs. In brief, the opportunity cost is the cost of choosing to use resources for one purpose that is measured by the sacrifice of the next best alternative for using these resources.
- **Accounting costs -** costs incurred on cash bases and only includes explicit costs.

D) Incremental cost and Sunk cost

Incremental cost-cost incurred due to a change in nature & the level of business. E.g. replacing machine, changing distribution channel.

Sunk cost- a cost that does not change due to nature and level of business. They are irrelevant for decision making for they do not vary in the production process. E.g. depreciation

Economic profit & Accounting profit

Profit refers to the difference between TR&TC of production of a firm. How ever, economists & accountants define profit differently.

Economists measure costs in terms of opportunity cost & they include both explicit and implicit costs in the calculation of profit.

i.e. Economic profit =TR-(implicit + explicit cost)

Accountants, on the other hand, do not take in to account implicit costs in the calculation of profit. To them TC of production includes only the sum of all explicit costs.

Accounting profit=TR-explicit cost

E) Depending on length of time, costs can be classified as:

i) Short run cost of production

ii) Long run cost of production

Short run cost of production: in the short run

-There is at least one variable and one fixed cost

-The time is too short for the producers to vary all the inputs

-Cost constitutes both fixed & variable costs (when they are added called total cost).

Fixed costs/ constant costs: They are parts of the total cost of the firm which does not vary with output, e.g.; expenditures on depreciation, rent of land and buildings, property taxes, etc.

Variable costs: These costs are directly dependent on the volume of output or service. Variable costs (e.g., expenditure on labor, raw material, etc.) increase but not necessarily in the same proportion as the increase in output.

From the TC we derive four types of costs:

ATC, AFC, AVC, & MC

TC=TFC+TVC, ATC=TC/Q, AFC=TFC/Q, AVC=TVC/Q, MC=△TC/△Q

Long run cost of production

-All inputs are varied, no fixed inputs.

-costs constitute only variable costs

-TC=TVC

Long run cost of production is useful to decide whether the optimum size of the plant is to be changed or not. Long run cost curves are derived from short run cost curves. Long run average cost curve is derived from short run average cost curves. Each point in the long run average cost curve corresponds to a point in a short run average cost curve, which is tangent to the long run at that point.

E.g. a firm takes a choice of three methods of production, each with different plant sizes; small, medium, large. These plants operate with the short run average costs of SAC_1, SAC_2, &SAC_3 respectively.

Productivity and costs

The shapes of the cost curves are mirror-image reflections of the shapes of the corresponding productivity curves. When one is increasing, the other is decreasing; when one is at minimum, the other is at a maximum.

Technical efficiency and economic efficiency

When choosing among existing technologies in the long run, firms are interested in the lowest cost & technically efficient methods of production. **Technical efficiency** in production means as **few inputs** as possible are used to produce a given output. **Economically efficient** method of production is the method that produces a given level of output at the **lowest possible cost.**

3.2.2. Determinants of the shape of the long-run cost curve

We saw that the law of diminishing marginal productivity accounted for the shape of the short-run average cost curve. The law of diminishing marginal productivity doesn't apply to the long run since in the long run all inputs are variable. The most important determinants of what is economically efficient in the long run are economies and diseconomies of scale.

3.2.2.1. Economies of scale

We say that production exhibits economies of scale when long-run average total costs decrease as output increases. For example, if producing 40,000 units costs a firm $16 million ($400 each), but producing 200,000 units costs the firm $40 million ($200 each), between 40,000 and 200,000 units, the production of these goods exhibits significant economies of scale. One can also say that there are increasing returns to scale.

Economies of scale account for the downward-sloping part. Cost per unit of output is decreasing.

3.2.2.2. Diseconomies of scale

We say that production exhibits diseconomies of scale when long-run average total costs increase as output increases. For example, if producing 200,000 units costs the firm $40 million ($200 each) and producing 400,000 units costs the firm $100 million ($250 each), there are diseconomies of scale associated with choosing to produce 400,000 rather than 200,000.

3.2.2.3. Constant returns to scale

Sometimes in a range of output a firm does not experience either economies of scale or diseconomies of scale. In this range there are **constant returns to scale** where long-run average total costs do not change with an increase in output. Constant returns to scale are shown by the

flat portion of the average total cost curve in Figure_ (b). Constant returns to scale occur when production techniques can be replicated again and again to increase output. This occurs before monitoring costs rise and team sprit is lost. In the long run, average costs initially fall because of economies of scale; then they are constant for a while, and finally they tend to rise due to diseconomies of scale.

Long-run production table (Table 3.2)

Quantity	Total costs of labor	Total costs of Machines	Total costs = TC_L+TC_M	Average total costs = $\frac{TC}{Q}$
11	$381	$254	$635	$58
12	390	260	650	54
13	402	268	670	52
14	420	280	700	50
15	450	300	750	50
16	480	320	800	50
17	510	340	850	50
18	549	366	915	51
19	600	400	1000	53
20	666	444	1100	56

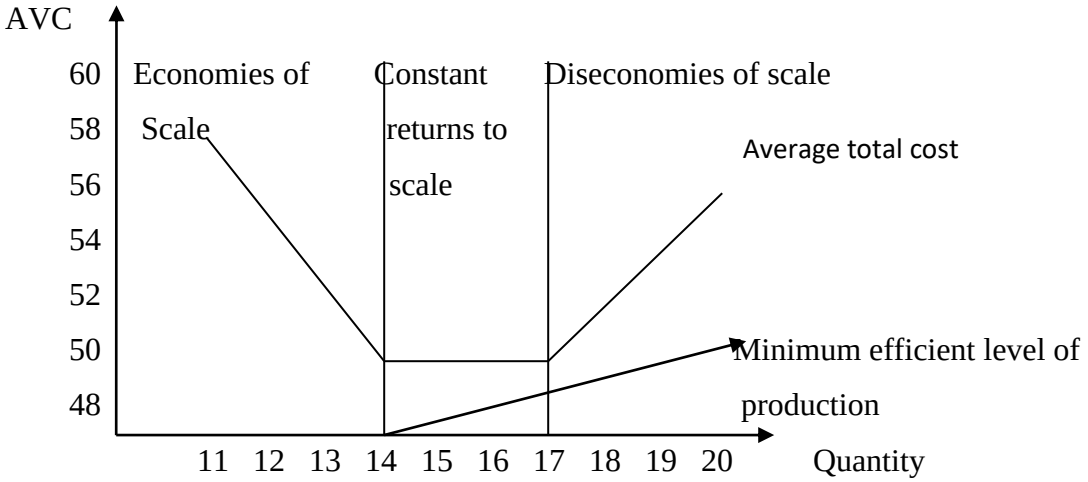


Figure 3.2. Long-run average cost curve

3.3. Production and Cost in Agricultural Firms (The relationship between production and costs)

A profit maximizing firm seeks to minimize its costs for a given output or to maximize its output for a given cost.

An isoquant curve is a curve showing all combinations of variable inputs that can be used to produce a fixed level of output. It is also known as equal-product or isoproduct curve.

If labor and capital were perfect substitutes (i.e.; zero complementarity), Isoquants would be linear. On the other hand, if labor and capital could not at all be substituted for each other (i.e.; zero substitutability) or perfect complementarity), the isoquant would take a rectangular shape.
Consider the production of good **A** which requires the use of two inputs (labor and capital). Suppose that it is possible to substitute labor for capital and capital for labor in the production process. Under this assumption, good **A** can be produced by using many different combinations of labor and capital.

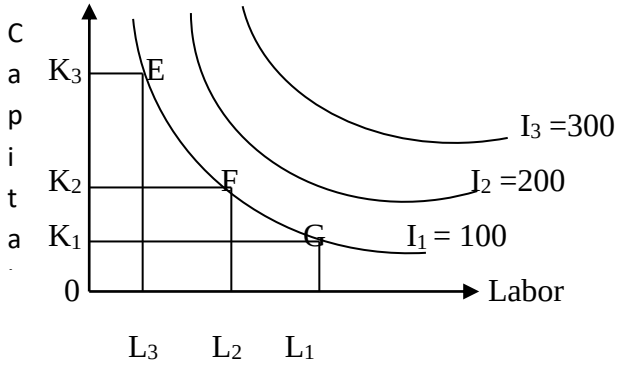


Figure 3.3. Isoquant map

To begin with, let us consider the information contained in Figure_. The figure shows three such isoquants. All the isoquants together constitute an isoquant map. Isoquant **I1**=100 indicates all possible combinations of labor & capital that can produce the same level of output.
The negative slope of an isoquant implies that if the producer reduces the amount of capital used, more labor has to be employed to remain on the same isoquant curve and vice versa. This means that the two inputs can be substituted for one another to maintain the same level of output. The rate at which one input is substituted for another keeping the level of output unchanged is known as marginal rate of technical substitution (MRTS). In our case,

$MRTS_{LK} = -\frac{dK}{dL}$ Where, $MRTS_{LK}$ is the marginal rate of technical substitution of labor for capital; dK is change in capital; dL is change in labor; since the slope of the isoquant $\left(\frac{dK}{dL}\right)$ is negative, the minus sign is added in order to make it positive.
When the amount of capital decreases, there is loss in output; when labor increases, there is gain in output. To remain on the same isoquant, the loss in output must be equal to the gain in output. The loss in output equals change in capital multiplied by the marginal product of capital (MP_K). The gain in output equals change in labor multiplied by the marginal product of labor (MP_L). Therefore,

$-dKMP_K = dLMP_L \qquad MRTS_{LK} = -\frac{dK}{dL} = \frac{MP_L}{MP_K}$

Any of the combinations on **I1** in Figure_ may be technically chosen to produce 100 units of good **A**, but not economically. Given input prices, the cost of production varies from point to point, and hence only one of the combinations on the isoquant **I1** gives minimum cost. The

question now is how to determine the input combination that leads to the minimum cost. This problem can be solved by combining the isoquant and isocost concepts.

An isocost curve shows all the different combinations of capital and labor which have the same cost. It is determined by the total outlay or total cost (TC) of the firm and input prices. Let us assume that the firm decides to incur total cost given by the following equation:

$TC = P_K K + P_L L$ Where, P_K *and* P_L are the unit prices of capital and labor respectively. From the above equation the amount of capital that can be bought equals,

$$K = \frac{TC}{P_K} - \frac{P_L}{P_K} L$$

And the amount of labor purchased equals

$$L = \frac{TC}{P_L} - \frac{P_K}{P_L} K$$

If the firm spends all of the total outlay on capital (no labor purchased), the amount of capital purchased is $K = \frac{TC}{P_K}$ which equals K_2 in Figure_. On the other hand, if the total outlay is spent on labor ($K = 0$), the amount of labor purchased is $L = \frac{TC}{P_L}$ and this equals L_2. The line connecting K_2 and L_2 gives an isocost line. Assuming constant P_L *and* P_K if the total outlay increases, the isocost curve shifts to the right. This means that more capital and labor are purchased. On the other hand, if the total outlay decreases (given input prices), the isocost curve shifts to the left. The absolute slope of an isoquant line is simply $\frac{P_L}{P_K}$.

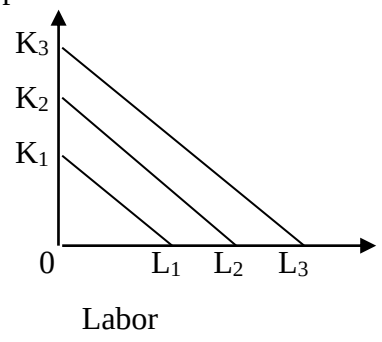


Figure 3.4. Isocost lines

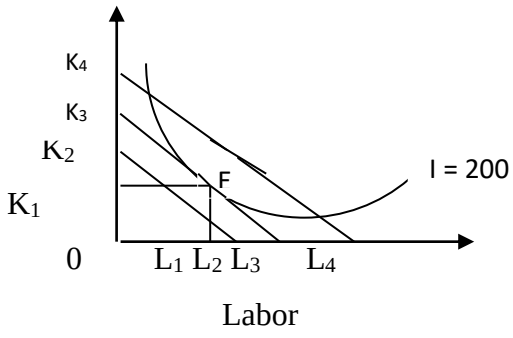


Figure 3.5. Least-cost combination of inputs

By bringing together on the same set of axis the firm's isoquants and isocosts; one can identify the least-cost level of output (point of producer equilibrium).

CHAPTER FOUR
THEORY OF FIRMS (MARKET STRUCTURE)

Pre tests

- Define the word market
- What kind of market structures do you know?

4.1 Perfectly competitive market:

Meaning and Definitions

Perfect competition is an extreme situation where competition reaches its maximum possible degree. The model of perfect competition however cannot be realistically expected to exist in totality in everyday life. It is highly theoretical, but the model provides a useful tool of economic analysis. Perfect competition is a theoretical market structure. It is primarily used as a benchmark against which other market structures are compared. The industry that best reflects perfect competition in real life is the agricultural industry.

In the short term, perfectly-competitive markets are not productively efficient as output will not occur where marginal cost is equal to average cost, but allocatively efficient, as output will always occur where marginal cost is equal to marginal revenue, and therefore where marginal cost equals average revenue. In the long term, such markets are both allocatively and productively efficient. Under perfect competition, any profit-maximizing producer faces a market price equal to its marginal cost. This implies that a factor's price equals the factor's marginal revenue product. This allows for derivation of the supply curve on which the neoclassical approach is based.

Characteristics of perfectly competitive market

This implies the number of buyers and sellers is so large that the share of each buyer (what each buyer purchases) in total demand and the share of each seller (what each seller offers for sale) in total supply is so small that no single buyer can affect the market price by changing his demand, not a single seller by changing his supply. This is likely to be true in many agricultural products.

Identical products: All the products offered by the individual suppliers must be exactly the same so that buyers will be completely indifferent as from which seller to buy.

Perfect knowledge: This means all participants (buyers and sellers) have knowledge about prices, quality, outputs levels and all other market conditions. Thus, no consumer would pay more than the market price and no supplier would be willing to sell for less than the market price.

Freedom entry and exit: There is no government interference in the form of price control, and no restriction which would prevent firms entering into the industry or leaving it. Under conditions of perfect competition, the interaction of the market supply and demand determines the market price for the industry, and each firm in the industry take the price as given. The firm under perfect competition is powerless to exert any influence on the price and is known as a ‘price taker’. Buyers are also price taker.

4.2 Pure Monopoly:

Meaning and Definitions

A monopoly exists when a specific person or enterprise is the only supplier of a particular commodity. (This contrasts with a monopsony which relates to a single entity's control of a market to purchase a good or service, and with oligopoly which consists of a few entities dominating an industry). Monopolies are thus characterized by a lack of economic competition to produce the good or service and a lack of viable substitute goods. The verb "monopolize" refers to the process by which a company gains the ability to raise prices or exclude competitors. In economics, a monopoly is a single seller. In law, a monopoly is business entity that has significant market power, that is, the power, to charge high prices. Although monopolies may be big businesses, size is not a characteristic of a monopoly. A small business may still have the power to raise prices in a small industry (or market).

A monopoly is distinguished from a monopsony, in which there is only one buyer of a product or service; a monopoly may also have monopsony control of a sector of a market. Likewise, a monopoly should be distinguished from a cartel (a form of oligopoly), in which several providers act together to coordinate services, prices or sale of goods. Monopolies, monopsonies and oligopolies are all situations such that one or a few of the entities have market power and therefore interact with their customers (monopoly), suppliers (monopsony) and the other companies (oligopoly) in a game theoretic manner – meaning that expectations about their behavior affects other players' choice of strategy and vice versa. This is to be contrasted with the model of perfect competition in which companies are "price takers" and do not have market power.

When not coerced legally to do otherwise, monopolies typically maximize their profit by producing fewer goods and selling them at higher prices than would be the case for perfect competition. Sometimes governments decide legally that a given company is a monopoly that doesn't serve the best interests of the market and/or consumers.

Characteristics of monopoly markets

A pure monopoly concerns the existence of a single seller of a good for which there is no close substitute.

- There is only one seller so that there is no distinction between firm and industry.
- The product is unique and there is no close substitute.
- The monopolist can exercise control over the price. The monopoly firm is, therefore, a price maker.
- The existence of barriers to entry in to the market. No other firm can supply the product because of legal, geographical or technological restrictions.
- There may or may not be advertisement.

Sources of monopoly market

The following are the main sources of monopoly:

Legal restrictions: in some cases, the government creates monopolies by law. It means that the government may establish or take over an industry and then prevents competition. It is illegal to compete with it. The Ethiopian Post service is an example of government monopoly.

Natural factor endowments: when one company or country is the sole or dominant owner of some kind of natural resource, it has an effective control over the supply of the resource and over the supply of any manufactured products derived from that resource.

High entry costs: in some industries the relatively high set-up costs create a formidable barrier to entry. In such cases, the existing monopoly firm, producing a large quantity output, benefits from economies of scale.

Patent: Governments encourage invention or innovation by granting exclusive production rights for a period of time for those who discover new products or new methods of production. As long as the patent is in effect the firm has a protected position and is a monopoly.

Efficiency: Effective management minimizes waste and leads to efficiency in production. This enables the firm to drive rivals out of business by undercutting price.

Agreements between producers: In addition to a sole supplier monopoly, it is also possible to find a number of firms acting together to create monopoly power. Such agreement is illegal in some countries. An example of international monopoly through collusion is Organization of Petroleum Exporting Countries (OPEC).

4.3 Monopolistic competition:

Meanings and Definitions

The real world is neither perfectly competitive nor perfectly monopolistic. The type of competition found in the real world lies between the two extremes. This form of competition is known as 'imperfect competition'. Monopolistic competition is one form of imperfect competition. It is a market structure in which many sellers compete to sell a differentiated product. As the name suggests, it combines some aspects of both perfect competition and monopoly.

Monopolistic competition is a type of imperfect competition such that competing producers sell products that are differentiated from one another as good but not perfect substitutes (such as from branding, quality, or location). In monopolistic competition, a firm takes the prices charged by its rivals as given and ignores the impact of its own prices on the prices of other firms.

In a monopolistically competitive market, firms can behave like monopolies in the short run, including by using market power to generate profit. In the long run, however, other firms enter the market and the benefits of differentiation decrease with competition; the market becomes more like a perfectly competitive one where firms cannot gain economic profit. In practice, however, if consumer rationality/innovativeness is low and heuristics are preferred, monopolistic

competition can fall into natural monopoly, even in the complete absence of government intervention.

In the presence of coercive government, monopolistic competition will fall into government-granted monopoly. Unlike perfect competition, the firm maintains spare capacity. Models of monopolistic competition are often used to model industries.

Characteristics of monopolistic competition

Many sellers and buyers: the market consists of large number of producers and customers. But the number may not be as large as that in perfect competition.

Product differentiation: Product differentiation is the production of different varieties or brands of the same product in the industry which consumers buy according to their preferences. Each supplier's product has unique qualities that cause some buyer as to prefer it to products of competing firms. The difference can be in quality, style, color, durability, brand name etc. The products of the many sellers are close substitutes for each other, but the individual seller has an element of monopolistic control over its own product.

Free entry and exit: in the monopolistically competitive markets, it is easy to set up new firms, and firms established in the industry find it easy to exit. No collusion: it is difficult for the firms to act in concert because they are many and they may not recognize their interdependence. This means the firms do not cooperate to fix prices so as to increase their group profits.

Because products are differentiated, the monopolistically competitive firm determines price like the monopolist. Price cuts increase quantity demanded by capturing part of the competitor's market. Unlike a perfectly competitive firm, a monopolistic competitive firm can raise its price and not loss all of its sales. In a monopolistic competition, some consumers will remain loyal to the firm's product and continue to purchase it even though the price has gone up and there are close substitutes.

Typical monopolistic competitors

Although it is difficult to identify industries that fit precisely our description of monopolistic competition, the following one approximate it: barbershops, grocery stores, laundries, clinics, retail stores, beauty shops, textile manufactures, furniture producers

I.4 Oligopoly market structure

Meanings and Definitions

The other type of market imperfection is oligopoly. Oligopoly is the form of market organization in which there are few sellers of a homogeneous or differentiated product. It is a market structure characterized by a small number of firms and a great deal of interdependence among them. Unlike monopolistic competition, oligopoly contains few firms that each oligopolist formulates its policies with an eye to their effect on rivals.

So any changes in the firm's price or output influence the sale and profits of competitors. Moreover, each firm must recognizes that changes in its own policies are likely to elicit changes in the policies competitor of this, as well. Because of interdependence oligopolists face a situation where the optimal decision of one firm depends on what other firms decide to do, and where the optimal decision of one firm depends on what other firms decide to do, and where there is opportunity, for both conflict and cooperation.

Oligopolistic competition can give rise to a wide range of different outcomes. In some situations, the firms may employ restrictive trade practices (collusion, market sharing etc.) to raise prices and restrict production in much the same way as a monopoly. Where there is a formal agreement for such collusion, this is known as a cartel. A primary example of such a cartel is OPEC which has a profound influence on the international price of oil.

Characteristics of oligopoly

Barriers to entry these refer to conditions that hinder the entry of new firms into the industry. These barriers may be in the form of ownership of resources, larger capital requirements, control of patents by existing firms. Few sellers 'few sellers' do not necessarily mean the number of sellers is very small. The number of firms may be large but few firms account for the major portion of the industry's productive capacity. That is, others are insignificant as far as their market share is concerned. Recognized mutual interdependence each firm recognizes that its profits are heavily dependent on the actions of the other firms in the industry.

Types of products under oligopoly products are either identical or differentiated. There is no other market structure that is more difficult to analyze than oligopoly. The difficulty arises from the interdependence nature of oligopolistic decisions. Oligopolists have a common interest in setting prices so that monopoly profits are made, and an individual interest in an attempting to get the maximum possible profit. These interests are contradictory in nature, and hence the economic theory of oligopoly concerning pricing and output decisions has not yet been standardized.

The sources of oligopoly

Economies of scale: scale economies may take it unprofitable for more than a few firms to coexist in the market

Patents or access: to a technology may exclude potential competitors

Huge capital investments and specialized inputs*:* are usually required to enter an oligopolies' industry and this act as natural barrier to entry.

Established firms*:* might have a loyal following of customers based on product quality and service that new firms may find very difficult to match.

The need to spend money for name recognition and market reputation may discourage entry

Oligopolistic Industries

Although it is common to analyze monopolistic competition and oligopoly separately, it is not always easy to draw very sharp distinctions between the two types of market structure. From the point of view of the number of producers and the ease of entry, some industries might be classified as monopolistically competitive. On the other hand, from the point of view of recognized mutual interdependence it might be reasonable to classify the same industries as oligopolistic, that is, they cannot be so easily classified. However, goods like oil, steel, tires, synthetic fibers, soap, and electric bulbs are usually produced under conditions of oligopoly.

CHAPTER FIVE
National Income Accounting

Whether economy of a country is growing or not is not known unless we can measure the total value of goods and services produced in the country for different years. This can be made with the help of national income accounting process.

'National Income Accounting' refers to the process of record keeping for the overall economic activities of a given country. It includes the goods and services produced in a country in (part of) a fiscal year.

5.1.The Basic Model: The Circular Flow Diagram

The economy consists of millions of people engaged in many activities-buying, selling, working, hiring, manufacturing, and so on. To understand how the economy works, we must find some way to simplify our thinking about all these activities. In other words, we need a model that explains, in general terms, how the economy is organized and how participants in the economy interact with one another. The circular-flow diagram offers a simple way of organizing all the economic transactions that occur between different sectors in the economy.

The circular-flow diagram is a visual model of the economy that shows how money flows through markets among the sectors of the economy. Accordingly, from the point of view of the number of sectors involved in the analysis or in the model, there are three major macroeconomic models:

a) Two sector model
b) Three sector model (closed economy model)
c) Four sector model (open economy model)

a) Two sector model

This model represents the case where there are only two sectors in the economy: the *household* sector and the *firm* sector. This model is represented by the relation or use of income by households i.e. consumers or households either consume or save their income. This relation is given by the following equation:

$$Y = C + S,$$

Where, Y = income,

C = Consumption expenditure,

S = Saving

Since saving is used for investment or saving is by itself a form of investment, S = I, the above equation can be rewritten as:

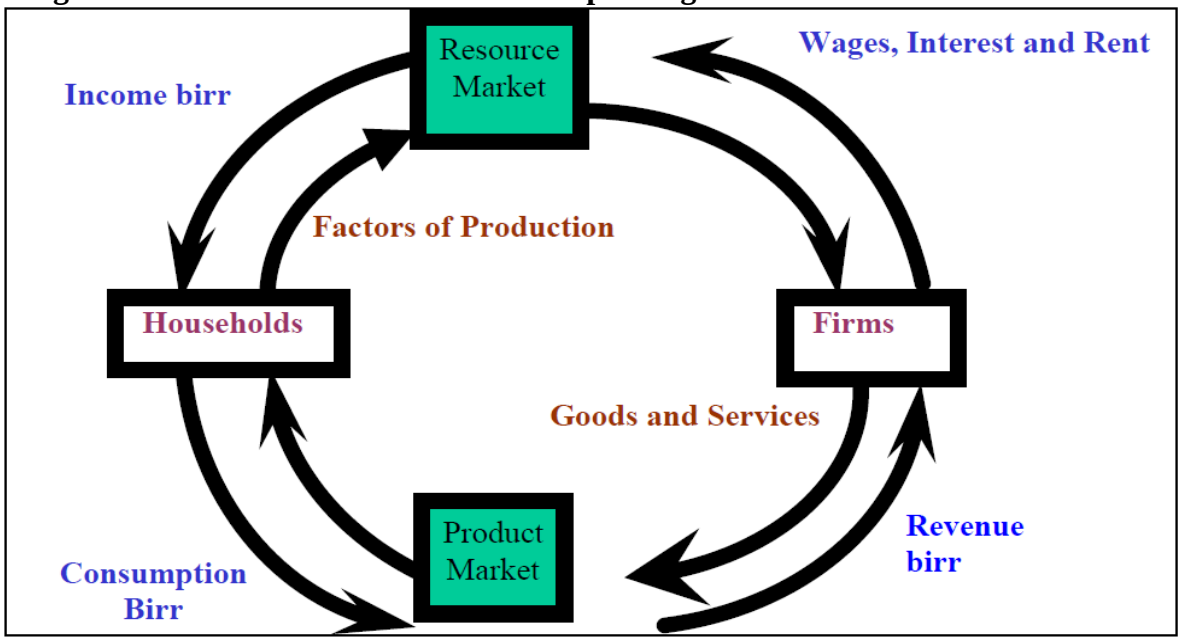
Y = C + I ,

Where, I = investment spending.

This model can be demonstrated using circular flow of income and expenditure as follows.

The inner loop of the circular-flow diagram represents the flows of goods and services, and factors of production between households and firms. Whereas the outer loop of the circular-flow diagram represents the corresponding monetary (Birr or dollar) flows.

Figure 5.1: Circular flow of income and spending in two sector model



Households sector is the owner of factors of production like land, labour and capital. These factors of payments are exchanged in resource market. In other words, factors of production are used by the firms and in return firms pay in terms of wages, interests and rent. These payments to the factors of production are nothing but the income (Y) from firms in the form of wage for their labour or in the form of rent on their land or interest in the form of capital. These factors of production spend part of their income on consumption goods (C) produced by firms and save the rest given by (S). For the sake of simplicity let us assume that all the income received by the households is consumed. We will introduce savings later. Business sectors produce the outputs to be sold in product market. These outputs are consumed by the households by spending their income. Business sector receives revenue from the consumption.

b) Three sector model (closed economy model)

This model incorporates government sector. Therefore, household, firms or the business sector and government are the three parties involved in the economy. This model is also known as

closed economy model because it does not include or consider trade with other countries. This model is represented by the following equation:

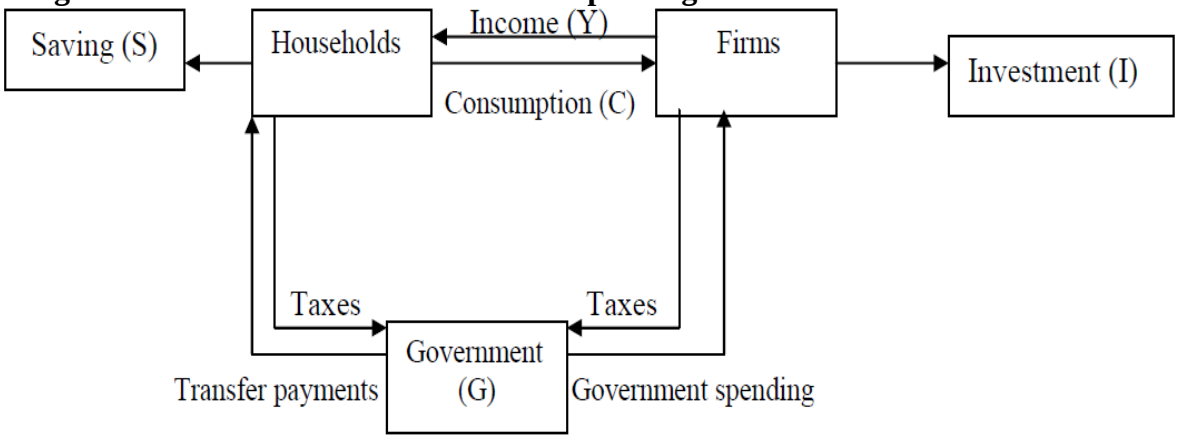
$Y = C + S + G$

$Y = C + I + G$,

Where, G = Government spending/ expenditure

This model can also be demonstrated using the use of circular flow of income and expenditure as follows (Figure 5.2). In addition to activities mentioned in two sector models above, households receive income from government transfer payments and pay tax to the government. In similar ways, business firms also sell their goods and services to the government and pay tax to the government. Government sector on their part use the tax income to finance its expenditure.

Figure 5.2: Circular flow of income and spending in three sector model



c) Four sector model (open economy model)

In addition to the three sector model, this model includes trade with other countries. As a result, the elements of trade such as import and export are incorporated in the model. This model is represented by the following equation:

$Y = C + I + G + NX$

Where: C= Consumption expenditure,

I= Investment expenditure,

G= Government expenditure

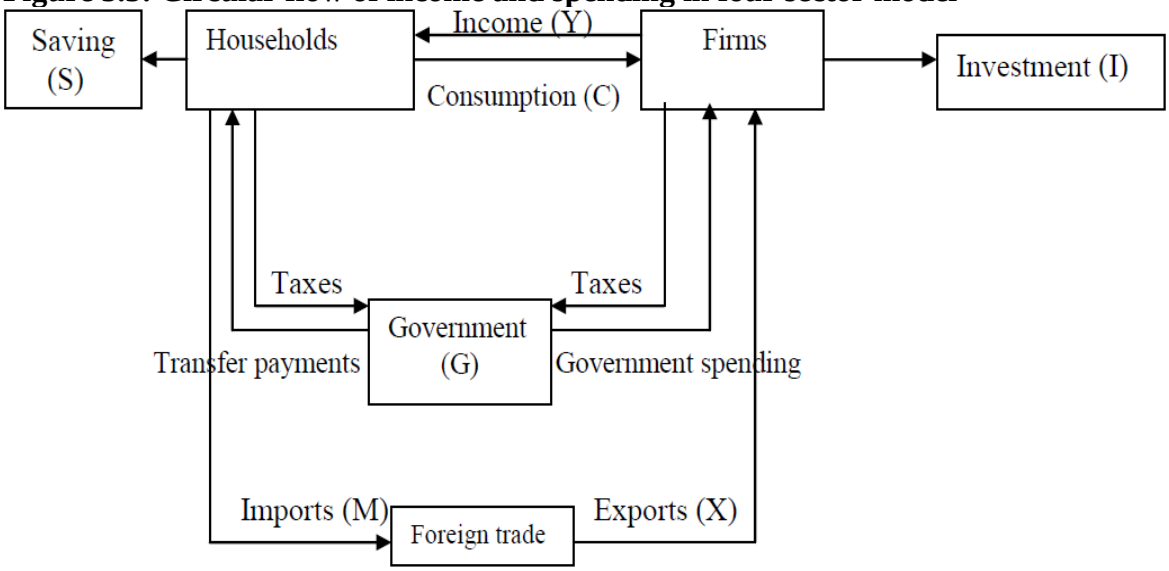
NX= Net export (X – M),

M= Import value,

X= export value

The four sector (open economy) model can also be demonstrated by the use of circular flow of income and expenditure as follows (figure 5.3).

Figure 5.3: Circular flow of income and spending in four sector model



By the introduction of foreign sector, the household sector and business sector can benefit by the imports and exports. Households can get imported products at the same time, business sector can also exports to other countries. Households spend on imported commodities by paying foreign exchange and business firms receive foreign exchange income by exporting their products. In this model the government plays a crucial role in regulating the foreign exchange market along with its previous role of collecting taxes and government spending.

Since almost all countries around the world have government involvement in the economy and have foreign trade with other countries, the third model is a more realistic one. Therefore, in this chapter, we discuss the national income accounting based on the four sector model or open economy model.

5.2. Real GDP versus Nominal GDP

Nominal GDP is the value of all final goods based on the prices existing during the time period of production. In other words, it is the price we pay in the market. Nominal GDP can grow in three ways:

- When output rises and prices remains constant.
- Prices rises and output remains unchanged.
- When both prices and output rises

The problem, then, is how to adjust GDP so that it reflects only changes in output and not changes in prices. This adjustment helps us in comparing the GDP over time when prices are changing. In order to know this we must understand the meaning of real GDP. Real GDP is the value of all final goods produced during a given time period based on the prices existing in a selected base year. The value of national output obtained by the use of such base year price is known as Real Gross Domestic Product (RGDP). Real GDP is also known as GDP in constant price or Birr. It is GDP adjusted for inflation.

Note that the base year may change from time to time. For instance a given base year price may be used only for five years and then another year price may be used as the base year price of the next five or more years.

Nominal gross domestic product (NGDP) is the relative magnitude of current GDP measured at current prices (current year) of goods and services. We can note this that only change in price level can lead to a change in NGDP even if there is no actual increase in output. As a result, most of the time, nominal GDP (NGDP) grows faster than real GDP (RGDP) mainly due to increases in prices of goods and services. This simply means that to get the most appropriate measure of national economic performances, nominal GDP (NGDP) should be adjusted to the real GDP (RGDP). This can be done by deflating NGDP when price is rising and by inflating it when price is falling. The adjustment factor is known as GDP deflator. GDP deflator is the ratio of nominal GDP to real GDP.

i.e. GDP deflator = $\left({}^{Nominal\ GDP}\!\!\diagup\!\!_{Real\ GDP} \right) = \left({}^{NGDP}\!\!\diagup\!\!_{RGDP} \right)$

Rearranging this equation, the value of real GDP is obtained by the following relation.

$$RGDP = \left({}^{NGDP}\!\!\diagup\!\!_{GDP\ deflator} \right)$$

Similarly, nominal GDP is calculated by multiplying real GDP by the GDP deflator.

$$\text{NGDP} = (RGDP)(GDP\ deflator)$$

The GDP deflator is an index measure of change in the general price level. GDP deflator can also be calculated from the outputs and services produced and their respective base year as well as from the current prices as follows.

$$GDP\ deflator = \frac{(Q_{C1}P_{C1}) + (Q_{C2}P_{C2}) + ...}{(Q_{C1}P_{B1}) + (Q_{C2}P_{B2}) + ...}$$

Where; Q_{Ci} = current unit of output or service or item 'i'. (i = 1, 2, 3, ……)

P_{Ci} = current price of output or service or item 'i' (i = 1, 2, 3, ……)

P_{Bi} = base year price of output or service or item 'i'(i = 1, 2, 3, ……)

Here, the numerator represents the value of the current year outputs (goods and services) at current price whereas the denominator represents the total value of the current output at base year price.

Illustrating Example: Given the following data/information on total output (goods and services) and price level (average price) in respective years, we can calculate nominal GDP, real GDP and GDP deflator. Note that in this particular exercise, we assume that all goods and services are measurable in similar units for the sake of simplicity. Let us take year 1970 as the base year and assume that the base year is changed to the year 1990. So the real GDP is calculated using price of the year 1970 up to the year 1989, but after the year 1989 we will use the 1990 price in calculating the RGDP, because year 1990 is selected as base year for the periods to come after this year. Thus, the result is given by the following table (Table 2.1).

From the table we can understand that in the year selected as base year both nominal GDP and real GDP are equal. This is because both values are calculated using the same price. That means, since the year which is selected as the base year is also the current year, the base year price and the current year price will be the same. From this example we can also see that nominal GDP may change simply because of change in price level even if there is no change in fiscal output. However, real GDP remains unchanged if there is no change in physical output. For instance, compare both values of real GDP and nominal GDP of the years 1980 and 1985. Since there was

no change in physical output (15 units in both years), there is no change in real GDP too, which remains 30 million Birr in both years. We can see the same case in the years 1994 and 1996 where there is no change in real GDP whereas the nominal GDP has increased simply because of an increase in price level (i.e. from 10 to 11 Birr per unit) even if there is no change in physical output which remains at 30 million Birr in both years.

Table 5.1: Real GDP and nominal GDP

Year	Unit of goods and services (in millions)	Price level	Nominal GDP (NGDP) (in million Birr)	Real GDP (RGDP) (in million Birr)	GDP deflator (NGDP/ RGDP)
1970	10	2	20	20	1
1975	12	3	36	24	1.5
1980	15	4	60	30	2
1985	15	6	90	30	3
1988	20	7	140	40	3.5
1990	25	8	200	200	1
1994	30	10	300	240	1.25
1996	30	11	330	240	1.375
1998	32	15	480	256	1.875
1999	34	18	612	272	2.25
2000	35	20	700	280	2.5

Note that consumer price index we have briefly explained in the previous section is calculated in the same manner as is the GDP deflator. That is the 'Consumer Price Index (CPI) is calculated by the relation:

$$\text{CPI}= \frac{(Q_{c1}P_{c1})+(Q_{c2}P_{c2})+.....}{(Q_{cl}P_{B1})+(Q_{c2}P_{B2})+.....}$$

Where; Q_{Ci} = current unit of consumption good or service or item 'i'. (i = 1, 2, 3, …)

P_{Ci} = current price of consumption good or service or item 'i'. (i = 1, 2, 3, ……)

P_{Bi} = base year price of consumption good or service or item 'i'. (i = 1, 2, 3, ….)

The difference between the two are that GDP deflator measures the prices of all goods and services (including expenditure of both firms and households) whereas CPI measures the prices of all goods and services purchased by consumers. The other difference between the two measures is that GDP deflator includes only outputs produced domestically whereas the CPI includes price of imported consumer goods as well.

5.3. Gross domestic product (GDP) and Gross national product (GNP)

In the national income accounting process, the terms 'Gross national product (GNP) and Gross Domestic Product (GDP) are usually used as an approximate of each other. But there is a distinction between gross domestic product (GDP) and gross national product (GNP). GDP is the value of final goods and services produced by domestically owned factors of production within a given period. But, GNP refers to the sum of values of goods and services produced by all citizens of a nation/a country all over the world in the country plus outside the country.

Gross domestic product (GDP) is the sum of values of goods and services produced in the country by citizens of the country and foreigners. Therefore, GDP is something related to territory of the country (i.e. GDP is territorial) whereas GNP is something related to citizenship (i.e. GNP is national). The difference between GNP and GDP equals to the net income earned by

foreigners (NFP). Let us take one example, Mr X is a citizen of Ethiopia and he is working in USA. Now the question is whether to include his income as a part of Ethiopian national income or not. Since he is not utilizing the resources of Ethiopia to earn, his income cannot be included in Ethiopian GDP but however it will be included in income of USA.

When GDP exceeds GNP, residents of a given country are earning less abroad than foreigners are earning in that country.

The relationship between the two values GDP and GNP is given as follows.
GNP – GDP = NFP, implying that; GNP = GDP + NFP,
Here, NFP is the net factor payment to citizens. This is nothing but the difference between income received by citizens of the country outside the country like Mr. X and the income received by foreigners in the country. Thus, GNP can be greater than or equal to or less than GDP depending on the value of the net factor payment. If the value of the net factor payment (NFP) is zero, it means both GDP and GNP are equal. If the value of net factor payment (NFP) is positive, then GNP is greater than GDP. This means, the output produced by citizens of the country is larger than what is produced in the territory of the country.

Example: Suppose Ethiopians abroad have produced output worth of 200 million Birr in the year 2005, and at the same time foreigners working in Ethiopia have produced output worth of 150 million Birr in the same year. If the Ethiopian GDP in that year is 800 million Birr, assuming that other things are constant, the net factor payment (NFP) and the GNP are given as follows:
NFP = GNP – GDP
NFP = income received by Ethiopians abroad *minus* income received by foreigners in Ethiopia.
NFP = 200 – 150 = 50 million Birr, then GNP can be calculated as follows;
GNP = GDP + NFP = 800 + 50 = 850 million birr
Most of the time, NFP is negligible, this means that most of the time the difference between GDP and GNP are found to be very small for several countries. Thus, GDP is used as a proxy of GNP.

5.4. Approaches to Measure GDP

A. The output approach

The output approach focuses on finding the total output of a nation by directly finding the total value of all goods and services a nation produces.

Because of the complication of the multiple stages in the production of a good or service, only the final value of a good or service is included in total output. This avoids an issue often called 'double counting, wherein the total value of a good is included several times in national output, by counting it repeatedly in several stages of production. In the example of meat production, the value of the good from the farm may be $10, then $30 from the butchers, and then $60 from the supermarket. The value that should be included in final national output should be $60, not the sum of all those numbers, $100. The values added at each stage of production over the previous stage are respectively $10, $20, and $30. Their sum gives an alternative way of calculating the value of final output.

Formulae: GDP (gross domestic product) at market price = value of output in an economy in a particular year - intermediate consumption

NNP at factor cost = GDP at market price - depreciation + NFIA *(net factor income from abroad)* - net indirect taxes

B. The expenditure approach

The expenditure approach is basically an output accounting method. It focuses on finding the total output of a nation by finding the total amount of money spent. This is acceptable, because like income, the total value of all goods is equal to the total amount of money spent on goods. The basic formula for domestic output combines all the different areas in which money is spent within the region, and then combining them to find the total output.

GDP = **C** + **I** + **G** + (**X** - **M**)

Where: **C** = household consumption expenditures / personal consumption expenditures
I = gross private domestic investment
G = government consumption and gross investment expenditures
X = gross exports of goods and services
M = gross imports of goods and services
Note: (**X** - **M**) is often written as $\mathbf{X_N}$, which stands for "net exports"

C. The Income Approach

The income approach equates the total output of a nation to the total factor income received by residents of the nation. The main types of factor income are:

- Employee compensation (= wages + cost of fringe benefits, including unemployment, health, and retirement benefits);
- Interest received net of interest paid;
- Rental income (mainly for the use of real estate) net of expenses of landlords;
- Royalties paid for the use of intellectual property and extractable natural resources.

All remaining value added generated by firms is called the *residual* or profit. If a firm has stockholders, they own the residual, some of which they receive as dividends. Profit includes the income of the entrepreneur- the businessman who combines factor inputs to produce a good or service.

NDP at factor cost = Compensation of employees + Net interest + Rental & royalty income + Profit of incorporated and unincorporated firms + Income from self-employment.

National income = NDP at factor cost + NFIA (net factor income from abroad)

GDP per capita (per person) is often used as a measure of a person's welfare. Countries with higher GDP may be more likely to also score highly on other measures of welfare, such a life expectancies. However, there are serious limitations to the usefulness of GDP as a measure of welfare:

- Measures of GDP typically exclude unpaid economic activity, most importantly domestic work such as childcare. This leads to distortions; for example, a paid nanny's income contributes to GDP, but an unpaid parent's time spent caring for children will not, even though they are both carrying out the same economic activity.
- GDP takes no account of the inputs used to produce the output. For example, if everyone worked for twice the number of hours, then GDP might roughly double, but this does not necessarily mean that workers are better off as they would have less leisure time. Similarly, the impact of economic activity on the environment is not measured in calculating GDP.
- Comparison of GDP from one country to another may be distorted by movements in exchange rates. Measuring national income at purchasing power parity may overcome this problem at the risk of overvaluing basic goods and services, for example subsistence farming.
- GDP does not measure factors that affect quality of life, such as the quality of the environment (as distinct from the input value) and security from crime. This leads to distortions - for example, spending on cleaning up an oil spill is included in GDP, but the negative impact of the spill on well-being (e.g. loss of clean beaches) is not measured.
- GDP is the mean (average) wealth rather than median (middle-point) wealth. Countries with a skewed income distribution may have a relatively high per-capita GDP while the majority of its citizens have a relatively low level of income, due to concentration of wealth in the hands of a small fraction of the population.

CHAPTER SIX
Macroeconomic Problems and Policies

Inflation and unemployment are the most important and recurrent economic problems that have characterized modern economic history throughout the world. The problem unemployment is associated with recessions and inflation is associated with the loss of purchasing power of our incomes. Most economic policy focuses on mitigating these, most serious, of problems in the macroeconomics.

6.1. Unemployment

6.1.1. Definitions and Concepts of Unemployment

Unemployment refers to a situation where workers of the working age could not find job while they are ready to work at prevailing market wage rate. The problem of unemployment and its intensity is usually *measured in terms of unemployment rate.* This is because talking about the actual number of unemployed people makes no sense for countries of different population size. Unemployment of 10,000 people in a given country with total population of 500,000 people is much compared to unemployment of 10,000 people in a country with total population of 500,000,000 people. Under this condition, we can say there is unemployment problem in the former where as there is no significant unemployment problem in the later. This is why it is better to measure unemployment problem in terms of unemployment rate.

Unemployment rate is defined as the percentage or the proportion of the labour force that is unemployed. **Labour force (L)** is the sum of both employed (E) and unemployed people of working age and working or ready to work (U). People who are employed may be either full time or part time employees. The labour force can be given by the following equation:

$$L = E + U \quad \text{---------- (5.1)}$$

Labour force does not include part of the working age population which are not working or looking for work, that is, the working age population that is not in the labour market. Retired persons, children, those who are either incapable of working or those who choose not to participate in the labor market are not counted in the labor force. The related concept is labour force participation rate. Labour force participation is the percentage of adult or working age population who are in the labour force. These two concepts are summarized by the following equations:

$$\textit{Unemployment Rate} = \left(\frac{Number\ of\ Unemployed}{Labour\ Force}\right) X\,(100)$$

$$\textit{Labour Force Participation Rate} = \left(\frac{Labour\ Force}{Adult\ Population}\right) X\,(100)$$

The concept of natural rate of unemployment is used as reference to the existence of unemployment problem. **Natural rate of unemployment** is the average rate of unemployment around which an economy fluctuates given by U/L = n at that point. It is related to the rate at which workers lose job (job loss rate) and the rate at which jobless workers find or get job (job finding rate). This can be more elaborated as follows using the steady state unemployment rate. **Steady state unemployment rate** is the point or condition in which the number of workers

leaving or losing job are equal to the number of unemployed getting or finding job given by the following equation:

$$fU=sE \text{ ---------- (5.2)}$$

Where, 'f' is rate of job finding,
's' is rate of job separation or loss
'U' is unemployed population,
'E' is Employed population

Rearranging equation (5.1) into E = L – U and substituting in (5.2), we obtain the following relation:

$fU = s(L – U)$ dividing both sides by $L \Rightarrow fU/L = (s/L)(L – U)$

$fU/L = s(1 – U/L)$ dividing both sides by 'f' we obtain

$U/L = s/(s + f) = n$ ---------- (5.3)

Where, U/L = n is the natural rate of unemployment

The policy implication of this is that any policy aimed at reducing the natural rate of unemployment should either reduce the rate of job separation (losing job), or increase the rate of job finding (f). Moreover, any policy that affects the rate of job separation (s) or job finding (f) also affects the natural rate of unemployment 'n'.

Examples:

1) If 90% of labour force of a given country with 10 million labour force population are employed on average, find the unemployment rate of the country.
2) If 3% of workers are on average leaving or losing their job and on average 40% of unemployed workers and workers newly joining labour market are finding or getting job, then what is the natural rate of unemployment?
3) In question number '1' if the size of the adult population of the country is 12 million, what is the labour force participation rate?

Solutions:

1. $L=E+U=90\%(L)+U \Rightarrow L-0.9L=U \Rightarrow U/L=0.1=10\%$
2. Given: job separation rate (s)=3%=0.03 and Job finding rate (f)=40%=0.4
   $U/L=n=(s/(s+f))=(0.03/(0.03+0.4))=0.0698=6.98\%$
3. Labour force participation rate=(labour force/adult population)x100
   = (10 million /12 million) x100=83.33%

6.1.2. Types of Unemployment

Depending on the *causes or sources* of the unemployment and the *duration* of the unemployment, we can divide unemployment into different categories. In this respect, economists divide unemployment into three or four major categories. These are frictional unemployment, structural unemployment, cyclical unemployment and/or seasonal unemployment.

I. Frictional unemployment

Frictional unemployment is a type of unemployment usually caused by constant changes in the labour market. It occurs due to two reasons. The first reason is when employers are not aware of the available workers and their job qualifications. The second reason is when workers are not fully aware of the jobs being offered. The basic cause of frictional employment is thus lack of

information flow among workers and employers called imperfect information. So for workers this is the period of unemployment until they get a job and the duration of unemployment is equal to the time it takes workers to search for a job. New graduating student from colleges and universities or training institutions are usually frictionally unemployed. This is the period when such people look for vacancies and apply to different offices getting interviewed and so on.

The change in composition of demand among economic sectors, industries or regions called **sectoral shift**. Due to this shift, the demand for workers also shift and some workers have to leave some sectors, industries or regions and look for jobs and join some other sectors which always involve frictional unemployment. The main characteristics of frictional unemployment are that:

- it affects large number and wide range of people
- it tends to be of short period
- certain amount of frictional unemployment is unavoidable

One of the *policy options* to solve such unemployment is *improving labor market information* (e.g. establishment of information office about workers and vacancies). Note that in trying to reduce frictional unemployment, some policies inadvertently increase the amount of frictional unemployment. One example of such policy is *unemployment insurance*. In this case, people will be reluctant in looking for job as soon as possible since they can collect some money because of their unemployment and prefer to stay for certain period after unemployment.

II. Structural Unemployment

Structural unemployment occurs due to the structural changes in the economy. These changes eliminate some jobs while they create some new jobs for people with new skill level. The skill sets take time to develop and hence some people lose their job simply because they do not have the new required skill(s). This problem arises from *mismatch* between the types of jobs that are available and type of job seekers. Such mismatch may be related to skill, education level, geographical area, age, etc. For instance, some skills may no longer be demanded. For instance, typing machines are replaced by computers. In this case type writers who do not have computer skill would lose their job and potentially become unemployed.

Some of the characteristics of this type of unemployment are that:

- It tends to be concentrated among certain group of people who are adversely affected by technological change.
- It tends to be long lasting (e.g. it takes time to the victims until they train themselves under new situation or new technology).

In this respect one of the *policy options* used to reduce such unemployment is *training workers* and improving labor mobility.

III. Cyclical Unemployment

Cyclical unemployment occurs due to general downturn in the business activities including production and demand for the products and services. During recession business conditions, only few goods are produced and for such low production, only few employment opportunities would be available. Employers are therefore, obliged to lay-off workers and cut back employment. As

we have tried to explain above, unemployment rate fluctuates around a line known as '*natural rate of unemployment'*. It is a rate where there is no cyclical unemployment or when all the unemployment is frictional and structural ones.

Cyclical unemployment is also known as **demand deficient unemployment**. Cyclical unemployment is the result of insufficient aggregate demand in the economy to generate enough jobs for those seeking them. It occurs during cyclical contraction of an economy (recession). The *policy instrument* to solve this problem is fiscal policy (for instance increasing government expenditure and reducing tax rates) and/or monetary policy (such as reducing interest rate and increasing money supply).

IV. Seasonal Unemployment

Seasonal unemployment is the type of unemployment that arises from a decline in the economic activity in *some seasons* (particular time in a year) and in *some sectors*. Therefore, seasonal unemployment results from fluctuations in demand for labor in these sectors and/or seasons. The special characteristic of this type of unemployment is that fluctuation can take a regular course of action and can be anticipated so that workers also make their own plan to move to particular sector in specific seasons to avoid such unemployment. For instance, workers can seek job in agricultural sector during the first season of cultivation and during harvest time. In other seasons the demand for labour in agriculture becomes low and as a result workers would look for job in other sectors.

In conclusion, with frictional and structural unemployment there could be enough jobs, but it is difficult to match job seekers with job vacancies. With cyclical unemployment, on the other hand, there are no enough jobs for job seekers. The duration of cyclical unemployment can be in between frictional and structural unemployment if the recession is handled and solved quickly (not as short duration as frictional and not as long duration as structural unemployment). Otherwise, it will take longer duration than both other types of unemployment.

Note that sometimes concepts such as underemployment, disguised unemployment and open unemployment are used to demonstrate different degrees of unemployment. *Underemployment* refers to the people who work below their capacity. For instance, a person may be employed for only 5 hours a day while he/she can and wants to work more than that. *Disguised unemployment* is the case where the worker is employed but adding nothing to the output. *Open unemployment* represents the formal definition we have discussed above where the person has no job at all while he/she is ready to work at the prevailing market wage rate.

6.2. Inflation

6.2.1. Concepts and Definition of Inflation

A birr today doesn't buy as much as it did ten years ago. The cost of almost everything may go up. This increase in the overall level of prices is called *inflation*, and it is one of the primary concerns of economists and policymakers.

The inflation rate measures how fast prices are rising. Inflation rate is the rate at which the average of these prices or overall price level increases from period to period. We can use the following formula to calculate inflation rate.

$$Inflation\ rate = \left(\frac{P_1 - P_0}{P_0}\right) X 100,$$

Where, P_0= previous year price index,

P_1=current year average price

If the inflation rate is very high, it is known as ***hyperinflation***. There is no consensus on when a particular rate of inflation becomes hyper but most of the economists would agree that inflation rate of about 100% per year would be hyper. Or it is defined as the level of inflation that exceeds 50% per month or a level of inflation which is greater than 1% per day. For example, if a person is having 200 birr in 1995 then the value of the same money is around 100 birr in 1996.

During hyperinflation, everything on the market becomes too costly. The purchasing power of money will fall down since one needs a larger amount of money to purchase small good or service. Thus, during hyperinflation money loses its role as store of value, unit of account and medium of exchange. Under this condition, the demand for money will be very low and people resort to using commodity money or bartering system where commodities are exchanged for commodity. In some countries people go wild by rioting and breaking the shops and stores in order to get food. Sometimes they even blame the government for the rise in price level leading to the toppling down of heads of the state.

6.2.2. Price indexes

Economists measure changes in the cost of living using the price indexes. Price indexes are the way we attempt to measure inflation and adjust aggregate economic data to account for price level variations.

There are three ways of measuring price index/ average price. These are: GDP deflator, consumer price index and producer price index.

a. GDP deflator: As we have just seen, nominal GDP reflects both the prices of goods and services and the quantities of goods and services the economy is producing. By contrast, by holding prices constant at base-year levels, real GDP reflects only the quantities produced. From these two statistics, we can compute a third, called the *GDP deflator,* which reflects the prices of goods and services but not the quantities produced.

The **GDP deflator** is calculated as follows:

$$GDP\ deflator = \left(Nominal\ GDP \Big/ Real\ GDP\right) x100$$

Because nominal GDP is current output valued at current prices and real GDP is current output valued at base-year prices, the GDP deflator reflects the current level of prices relative to the level of prices in the base year. Because nominal GDP and real GDP must be the same in the base year, the GDP deflator for the base year always equals 100.

Imagine that the quantities produced in the economy rise over time but prices remain the same. In this case, both nominal and real GDP rise together, so the GDP deflator is constant. Now suppose, instead, that prices rise over time but the quantities produced stay the same. In this

second case, nominal GDP rises but real GDP remains the same, so the GDP deflator rises as well. In both cases, the GDP deflator reflects what's happening to prices, not quantities.

Numerical example: Suppose that for year 2001, nominal GDP is dollar 200, and real GDP is dollar 200, so the GDP deflator is 100. Again assume that for the year 2002, nominal GDP is dollar 600, and real GDP is dollar 350, so the GDP deflator is 171. Because the GDP deflator rose in year 2002 from 100 to 171, we can say that the price level increased by 71 percent.

b. Consumer price index (CPI): is a measure of the overall cost of the goods and services bought by a typical consumer. In other words, CPI measures the cost of buying a fixed basket of goods and services representative of the purchases of urban consumers.

$$CPI = \left(\frac{Total\,\text{exp}enditure\,on\,market\,basket\,in\,current\,year}{Total\,\text{exp}enditure\,on\,market\,basket\,in\,base\,year} \right) * 100$$

To calculate the consumer price index and the inflation rate, data on the prices of thousands of goods and services is required. However, to see exactly how these statistics are constructed let's consider a simple economy in which consumers buy only two goods-bread and apple. Table 6.1 shows the *five steps* that we follow.

1. *Fix the Basket.* Set by surveying consumers and finding the basket of goods and services that the typical consumer buys. In the example in the table, the typical consumer buys a basket of 4 breads and 2 apples.
2. *Find the Prices.* The table shows the prices of bread and apple for three different years.
3. *Compute the Basket's Cost.* The table shows the calculation for each of the three years.
4. *Choose a Base Year and Compute the Index.* To calculate the index, the price of the basket of goods and services in each year is divided by the price of the basket in the base year, and this ratio is then multiplied by 100. In the example in the table, the year 2001 is the base year and hence, the consumer price index is 100. The consumer price index is 175 in 2002. This means a basket of goods that costs dollar 100 in the base year costs dollar 175 in 2002.
5. *Compute the Inflation Rate.* Finally, use the consumer price index to calculate the **inflation rate,** which is the percentage change in the price index from the preceding period. That is, the inflation rate between two consecutive years is computed as follows:

$$\text{Inflation rate in year 2} = \left(\frac{\text{CPI in year 2} - \text{CPI in year 1}}{\text{CPI in year 1}} \right)$$

In our example, the inflation rate is 75 percent in 2002 and 43 percent in 2003.

Table 6.1: Calculating the consumer price index and the inflation rate

Step 1: Survey Consumers to Determine a Fixed Basket of Goods		
basket of 4 bread and 2 apples		
Step 2: Find the Price of each Good in each Year		
Year	Price of Bread	Price of Apple
2001	dollar 1	dollar 2
2002	dollar 2	dollar 3
2003	dollar 3	dollar 4
Step 3: Compute the Cost of the Basket of Goods in each Year		
2001	(dollar1 per bread x 4 bread) + (dollar2 per apple x 2 apples) = dollar8	
2002	(dollar2 per bread x 4 bread) + (dollar3 per apple x 2 apples) = dollar14	
2003	(dollar3 per bread x 4 bread) + (dollar4 per apple x 2 apples) = dollar20	
Step 4: Chose One Year as a Base Year (2001) and Compute the Consumer Price Index		
2001	(dollar8/dollar8) x100=100	
2002	(dollar14/dollar8)x100=175	
2003	(dollar20/dollar8)x100=250	
Step 5: Use the Consumer Price Index to Compute the Inflation Rate from Previous Year		
2002	(175-100)/100 x100 =75%	
2003	(250-175)/175 x100 =43%	

The GDP Deflator versus the Consumer Price Index

Economists and policymakers monitor both the GDP deflator and the consumer price index to gauge how quickly prices are rising. Usually, these two statistics tell a similar story. Yet there are three important differences that can cause them to diverge. The following are the three main differences between CPI and GDP deflator:

- The deflator measures the prices of a much wider group of goods than the CPI does.
- The CPI measures the cost of a given basket of goods, which is the same from year to year. The basket of goods included in the GDP deflator, however, differs from year to year, depending on what is produced in the economy in each year. When corn crops are large, corn receives a relatively large weight in the computation of the GDP deflator.
- The CPI directly includes prices of imports, whereas the deflator includes only prices of goods produced in the country, say Ethiopia.

c. Producer price index (PPI): measures the cost of a basket of goods and services bought by firms rather than consumers. Because firms eventually pass on their costs to consumers in the form of higher consumer prices, changes in the producer price index are often thought to be useful in predicting changes in the consumer price index.

$$\text{PPI} = \left(\frac{\text{Cost of production in current year}}{\text{Cost of production in base year}} \right)$$

6.2.3. Causes and Effects of inflation

6.2.3.1. Causes of inflation

Some of the major causes or sources of inflation are an increase in money supply, high demand for goods and services and shortage of supply of goods and services.

There are three theories of inflation that arise from the *real* conduct of the macroeconomics. These three theories are demand-pull, cost-push, and pure inflation. There is also a fourth theory that suggests that inflation has little or nothing to do with the *real output* of the economy, this is called the quantity theory of money.

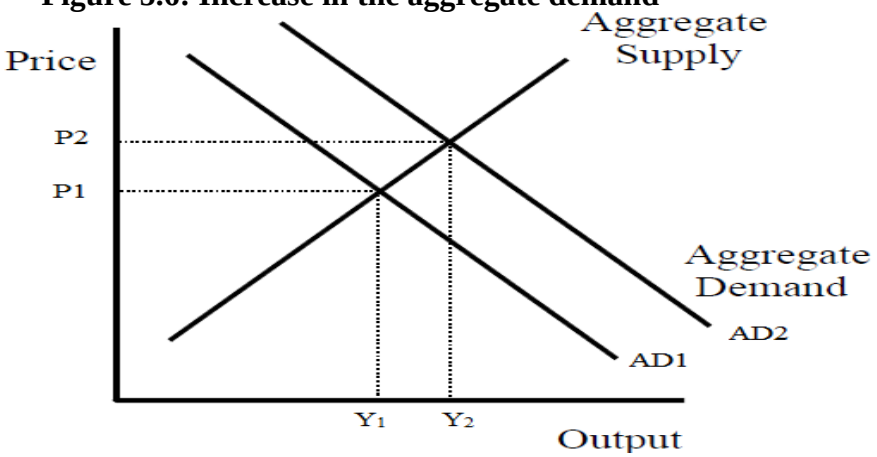
a. Demand - Pull Inflation

When there is high or increase in demand for goods and services individuals or users of the goods and services would be ready to pay higher price. Suppliers have the incentive to charge higher price when there is high demand for their product. As a result of these incentives, the price level will swing upward. This type of inflation is known as **demand pull inflation**.

On aggregate demand-aggregate supply model, as the aggregate demand shifts to the right or increases from (AD_1 to AD_2 see the following figure), all prices increase (P_1 to P_2). However, this increase in aggregate-demand is also associated with an increase in total output (Y_1 to Y_2). Total output is associated with employment (remember Okun's Law). In other words, even though this increase in aggregate demand causes inflation, it does not result in lost output, hence unemployment.

Policy measures designed to control demand-pull inflation, will shift the aggregate demand curve to the left, (i.e., reduce aggregate demand) and this reduction in aggregate demand is associated with loss of output, hence increased unemployment.

Figure 5.6: Increase in the aggregate demand

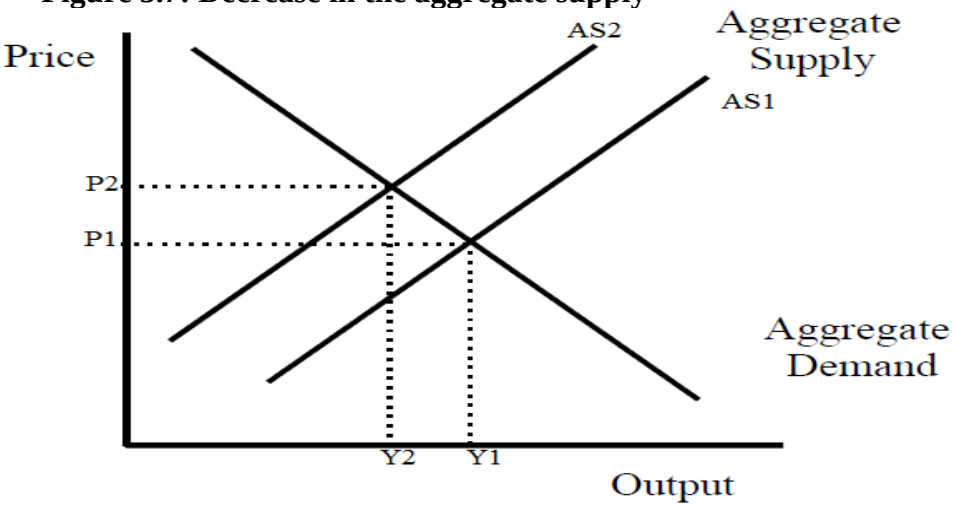


b. Cost - Push Inflation

When there is high cost of production, there will be shortage of supply of goods and services. Then to cover their high cost of production and get profit margin producers of goods and services charge higher prices. Moreover, when there is shortage of supply, consumers or users of goods and services are willing to pay higher price to get the limited goods and services. This means, both actions lead to an increase in price level. Such inflation or an increase in price level arises from the supply side or from the cost of production and is therefore, called **cost push inflation**. For instance, there was a supply shock of oil during 1970s across the world and this increased cost of production and distribution that finally resulted in inflation during the period. If the problem of inflation arises along with low aggregate output or during economic stagnation or recession the situation is called **stagflation**.

Again using an aggregate supply-aggregate demand approach, cost-push inflation results from a decrease in the aggregate supply curve. The following diagram shows a shift to the left or decrease in aggregate supply curve (from AS_1 to AS_2) and a price increase from P_1 to P_2.

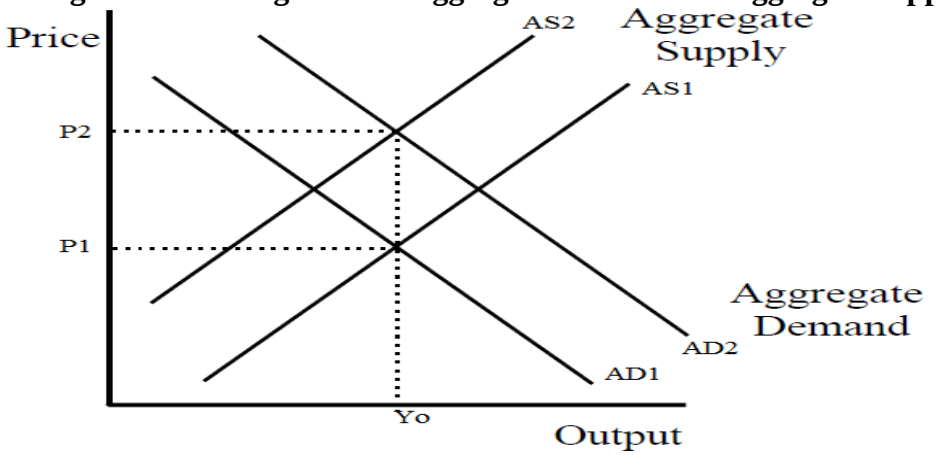
Figure 5.7: Decrease in the aggregate supply



c. Pure Inflation

Pure inflation results from an increase in aggregate demand and a simultaneous decrease in aggregate supply. For output to remain unaffected by these shifts in aggregate demand and aggregate supply, then the increase in aggregate demand must be exactly offset by an equal decrease in aggregate supply.

Figure 5.8: Change in both aggregate demand and aggregate supply



Notice in this diagram, that aggregate supply shifted to the left, or decreased (from AS_1 to AS_2) by exactly the same amount that the aggregate demand curve shifted to the right or increased (from AD_1 to AD_2). The result is that output remains exactly the same, but the price level increased.

d. Quantity Theory of Money

Inflation, in the monetarist view, can only occur if the money supply is increased which permits all prices to increase. If the money supply is not increased there can be changes in relative prices, for example, oil prices can go up, but there has to be offsetting decreases in the prices of other

commodities. An increase in all prices or in the price of a particular good, therefore, is a failure of the central bank (Fed) to appropriately manage the money supply.

6.2.3.2. Effects/consequences of Inflation

Whatever its sources or its causes, inflation has some common major effects on the society or the country where inflation occurred. Some of these effects are high living costs, higher wage rates, excess nominal money supply over the real products, shortage of supply of goods and services due to high cost of production, very low value of domestic currency (or money), expensive imports, and so on.

When inflation occurs, each unit of consumer goods is bought at higher price and living expenses become higher unless the income of individuals increases as well. This is why workers push wage rate upward during inflation period. This implies that the purchasing capacity of money or domestic currency deteriorates. Compared to foreign currencies the value of domestic currency becomes very low and becomes subject to be changed to smaller amount of foreign currency. Thus, when one purchases smaller amount of foreign commodity or imports becomes expensive. There will be inflow of foreign currency if exports and imports are elastic. But normally, in developing countries petroleum products constitute the major part of imports and hence imports are inelastic. Imports remain the same despite it becomes costly for the domestic country, the country becomes poorer.

The effects of inflation may impact different people in different ways. Creditors and those living on a fixed income will generally suffer. However, debtors and those whose incomes can be adjusted to reflect the higher prices will not, and perhaps this group may even benefit from higher rates of inflation.

If inflation is *fully anticipated* and people can adjust their nominal income or their purchasing behavior to account for inflation then there will likely be no adverse effects, however, if people cannot adjust their nominal income or consumption patterns people will likely experience adverse effects. This is the same as if people experience unanticipated inflation. Normally, if you cannot adjust income, are a creditor with a fixed rate of interest or are living on a fixed income you will pay higher prices. The result is that those individuals will see their standard of living eroded by inflation.

Debtors, whose loans specify a fixed rate of interest, typically benefit from inflation because they can pay loans-off in the future with money that is worth less. It is this paying of loans with money that purchases less that harms creditors.

Savers may also find themselves in the same position as creditors. If savings are placed in long-term savings certificates that have a fixed rate of interest, inflation can erode the earnings on those savings substantially. Savers that anticipate inflation will seek assets that vary with the price level, rather than risk the loss associated with inflation.

Inflation will affect savings behavior in another way. If a person fully anticipates inflation, rather than to save money now, consumers may acquire significant debt at fixed interest rates to take advantage of the potential inflationary leverage caused by fixed rates. Rather than to save now,

consumers spend now. Therefore, inflation typically creates expectations among people of increasing prices, and if people increase their purchases aggregate demand will increase. An increase in aggregate demand will cause demand-pull inflation. Therefore, inflationary expectations can create a spiraling of increased aggregate-demand and inflationary expectations that can feed off one another. At the other extreme, recessionary expectations may cause people to save, that results in reduced aggregate demand, and another spiral effect can result (but downwards).

6.3. Macroeconomics Policy

6.3.1. Fiscal Policy

Fiscal policy is the use of government taxes and spending to alter macroeconomic outcomes. Fiscal policy refers to deliberate changes in tax or spending legislation.

The three stances of fiscal policy are the following:

 Neutral fiscal policy is usually undertaken when an economy is in neither a recession nor a boom. The amount of government deficit spending (the excess not financed by tax revenue) is roughly the same as it has been on average over time, so no changes to it are occurring that would have an effect on the level of economic activity.

 Expansionary fiscal policy is used by the government when trying to balance the contraction phase in the business cycle. It involves government spending exceeding tax revenue by more than it has tended to, and is usually undertaken during recessions. Examples of expansionary fiscal policy measures include increased government spending on public works (e.g., building schools) and providing the residents of the economy with tax cuts to increase their purchasing power (in order to fix a decrease in the demand).

 Contractionary fiscal policy, on the other hand, is a measure to increase tax rates and decrease government spending. It occurs when government deficit spending is lower than usual. This has the potential to slow economic growth if inflation, which was caused by a significant increase in aggregate demand and the supply of money, is excessive. By reducing the economy's amount of aggregate income, the available amount for consumers to spend is also reduced. So, contractionary fiscal policy measures are employed when unsustainable growth takes place, leading to inflation, high prices of investment, recession and unemployment above the "healthy" level of 3%-4%.

The Keynesian view of economics suggests that increasing government spending and decreasing the rate of taxes are the best ways to have an influence an aggregate demand, stimulate it, while decreasing spending and increasing taxes after the economic boom has already taken place. Additionally, Keynesians argue that expansionary fiscal policy should be used in times of recession or low economic activity as an essential tool for building the framework for strong economic growth and working towards full employment. In theory, the resulting deficits would be paid for by an expanded economy during the boom that would follow; this was the reasoning behind the New Deal. Governments can use a budget surplus to do two things: to slow the pace of strong economic growth; to stabilize prices when inflation is too high.

6.3.2. Monetary Policy

Monetary policy is the macroeconomic policy laid down by the central bank. It involves management of money supply and interest rate used by the government of a country to achieve macroeconomic objectives like inflation, consumption, growth and liquidity.

Monetary policies are implemented through different tools, including the adjustment of the interest rates, purchase or sale of government securities, and changing the amount of cash circulating in the economy. The central bank or a similar regulatory organization is responsible for formulating monetary policies.

Expansionary Monetary Policy: It is a monetary policy that aims to increase the money supply in the economy by decreasing interest rates, purchasing government securities by central banks, and lowering the reserve requirements for banks. An expansionary policy lowers unemployment and stimulates business activities and consumer spending. The overall goal of the expansionary monetary policy is to fuel economic growth. However, it can also possibly lead to higher inflation.

Monetary policies can influence the level of unemployment in the economy. For example, an expansionary monetary policy generally decreases unemployment because the higher money supply stimulates business activities that lead to the expansion of the job market.

Contractionary Monetary Policy: The goal of a contractionary monetary policy is to decrease the money supply in the economy. It can be achieved by raising interest rates, selling government bonds, and increasing the reserve requirements for banks. The contractionary policy is utilized when the government wants to control inflation levels.

Functions of money: Basically, there are three major functions of money. These are:

1. ***Money as a Medium of Exchange***: - Goods and services are sold in exchange for money and bought in exchange with money. In short, money increases trade by providing a much more convenient method of exchange than a difficult barter system.

2. ***Money as a Unit of Account****:* - How does one know the GDP value of a county? It is simply the money value of the goods and services produced in a country in a given year. In this case, money is used as a unit of account that is as base for measure of values of goods and services. The unit of account is a function of money to provide common measurement of the relative value of goods and service. With this function, we can easily compare one kg of *teff* with three loaves of bread if their money values are the same. In this comparison, money serves as a unit of account.

3. ***Money as Store of Value****:* - The store of value is the ability of money to hold value over time. You can keep your money in a safe place for a year or years and use it in the future. Values of goods and services produced can be changed to money and stored or kept in the form of money. This is because keeping the products for future sale may not be possible or may be wasteful or costly. However, high inflation or hyperinflation generally affects the functions of money as store of value and medium of exchange.