# NATIONAL OPEN UNIVERSITY OF NIGERIA

## FACULTY OF SOCIAL SCIENCES

## COURSE CODE: ECO 431

## COURSE TITLE: ADVANCED MICRO-ECONOMICS

**COURSE GUIDE**

# ECO 431
# ADVANCED MICROECONOMICS

**Course Team** Dr. S. O. Akinleye (Course Developer) – UNILAG
Dr. (Mrs.) G. A. Adesina-Uthman (Course Editor) – NOUN

**Course Reviewer:** Professor S.J. Mohammed (A.B.U, Zaria)

NATIONAL OPEN UNIVERSITY OF NIGERIA

National Open University of Nigeria
Headquarters
14/16 Ahmadu Bello Way
Victoria Island, Lagos

Abuja Office
5 Dar es Salaam Street
Off Aminu Kano Crescent
Wuse II, Abuja

e-mail: centralinfo@nou.edu.ng
URL: www.nou.edu.ng

Published by
National Open University of Nigeria

Printed 2014
Reviewed 2023

ISBN: 978-058-855-9

**CONTENTS** **PAGE**

## Introduction

**ECO 431-Advanced Microeconomics** is a core course, which carries two credit units for final year students of Economics in the School of Arts and Social Sciences at the National Open University of Nigeria. ECO 431 is a very useful course to you in your academic pursuit as it would help you to gain an in-depth understanding on how the economy is operated from the individual units. This course guide tells you what principles of economics (microeconomic theory) entails, what course materials you will be using and how you can work your way through these materials. It suggests some general guidelines for the amount of time required of you on each unit in order to achieve the course aims and objectives successfully. It also provides you some guidance on your Tutor-Marked Assignments (TMAs) as contained herein.

## What You Will Learn in this Course

You will be taught the introduction to the discipline of economics and be provided a basic understanding of how individual unit of the economy work together to determine the economic activities of a nation.

## Course Content

This course builds on the foundation of elementary economics on studying the effects of individual unit behavior on the economy. Topics covered include: Market Structure, Comparative Static and Dynamic Analysis, Theory of Demand, Consumer Behavior, Optimization, Utility Maximization, Welfare Economics, General Equilibrium, Linear Programming, etc.

## Course Aims

There are 16 study units in this course and each unit has its objectives. You should read the objectives of each unit and bear them in mind as you go through the unit. In addition to the objectives of each unit, the overall aims of this course are to:

- introduce you to the basic analysis viz: comparative static and dynamic analysis
- teach you the concept of linear programming
- expose you to the analysis of utility maximization
- expose you extensively to market structure

- enlighten you on the theory of cost and production.

## Course Objectives

On successful completion of the course, you should be able to:

- define both functions of the economy and the power and breadth of economics
- explain consumer utility and consumer surplus
- describe the concept of linear programming and welfare economics.

## Working Through This Course

To successfully complete this course, you are required to read the study units, referenced books and other materials on the course. Each unit contains Self-Assessment Exercise (SAE). At some points in the course, you may be required to submit assignments for assessment purposes. At the end of the course, there is a final examination.

## Course Materials

The major components of the course, what you have to do and how you should allocate your time to each unit in order to complete the course successfully on time are listed as follows:

1. Course Guide
2. Study Unit
3. Textbook and References
4. Assignment File
5. Presentation Schedule

## Study Units

There are five modules in this course broken into16 study units as follows:

**Module 1 Introductory Microeconomics**

Unit 1 Introduction to Microeconomics
Unit 2 Economic Systems and Decision Making
Unit 3 Comparative Static and Dynamic Analysis

**Module 2 Theory of Demand, Consumer Behavior and Optimization**

## Textbooks and References

Each unit contains a list of materials for further reading. Try to get as many as possible of those textbooks. They are meant to deepen your knowledge of the course.

## Assignment File

In this file, you will find all the details of the work you must submit to your tutor for marking. The marks you obtain from these assignments will count towards the final mark you obtain for this course. Further information on assignments will be found in the Assignment File itself and later in this Course Guide in the section on assessment.

## Presentation Schedule

The Presentation Schedule included in your course materials gives you the important dates for the completion of TMAs and attending tutorials. Remember, you are required to submit all your assignments by the due date. You should guard against falling behind in your work.

## Assessment

Your assessment will be based on TMAs and a final examination which you will write at the end of the course.

## Tutor-Marked Assignment (TMA)

There is one Tutor-Marked Assignment in each unit of this course. You will submit all the assignments. You are encouraged to work all the questions thoroughly. The TMAs constitute 30% of the total score.

You will be able to complete your assignments from the information and materials contained in your text books and study units. However, it is desirable that you demonstrate that you have read and researched more widely than the required minimum. You should use other references to have a broad viewpoint of the subject and also to give you a deeper understanding of the subject.

When you have completed each assignment, send it, together with a TMA form, to your tutor. Make sure that each assignment reaches your tutor on or before the deadline given in the Presentation File. If for any reason, you cannot complete your work on time, contact your tutor before the assignment is due to discuss the possibility of an extension. Extensions will not be granted after the due date unless there are exceptional circumstances.

## Final Examination and Grading

The final examination for ECO 431 will be of three hours' duration and have a value of 70% of the total course grade. The examination will consist of questions which reflect the types of Self-Assessment Exercises and Tutor-Marked Assignments you have previously encountered. All areas of the course will be assessed.

Use the time; between finishing the last unit and sitting for the examination to revise the entire course material. You might find it useful to review your Self-Assessment Exercises, TMAs and comment on them before the examination. The final examination covers information from all parts of the course.

## Course Marking Scheme

The scheme presented below indicates the total marks (100%) allocation.

**Assessment Marks**

| | |
|---|---|
| Assignment | 30% |
| Final Examination | 70% |

**Total 100%**

## How to get the Most from this Course

In distance learning, the study units replace the university lecturer. This is one of the great advantages of distance learning; you can read and work through specially designed study materials at your own pace and at a time and place that suit you best. Think of it as reading the lecture instead of listening to a lecturer. In the same way that a lecturer might set you some reading to do, the study units tell you when to read your books or other materials, and when to embark on discussion with your colleagues. Just as a lecturer might give you an in-class exercise, your study units provide exercises for you to do at appropriate points.

Each of the study units follows a common format. The first item is an introduction to the subject matter of the unit and how a particular unit is integrated with the other units and the course as a whole. Next is a set of learning objectives. These objectives let you know what you should be able to do by the time you have completed the unit.

You should use these objectives to guide your study. When you have finished the unit, you should go back and check whether you have achieved the objectives. If you make a habit of doing this, you will significantly improve your chances of passing the course and getting the best grade.

The main body of the unit guides you through the required reading from other sources. This will usually be either from your text books or from a reading section. Some units require you to undertake practical overview of historical events. You will be directed when you need to embark on discussion and guided through the tasks you must do.

The purpose of the practical overview of some certain historical economic issues are in two folds. First, it will enhance your understanding of the material in the unit. Second, it will give you practical experience and skills to evaluate economic arguments, and understand the roles of history in guiding current economic policies and debates outside your studies. In any event, most of the critical thinking skills you will develop during studying are applicable in normal working practice, so it is important that you encounter them during your studies.

Self-Assessment Exercises are interspersed throughout the units, and ways to answer them are already integrated in the units. Working through these tests will help you to achieve the objectives of the unit and prepare you for the assignments and the examination. You should do each Self-Assessment Exercise as you come into the study unit. Also, ensure to master some major historical dates and events during the course of studying the material.

The following is a practical strategy for working through the course. If you run into any trouble, consult your tutor. Remember that your tutor's job is to help you. When you need help, do not hesitate to call and ask your tutor to provide it.

## Read this Course Guide Thoroughly

Organise a study schedule. Refer to the `Course Overview' for more details. Note the time you are expected to spend on each unit and how the assignments relate to the units. Important information, e.g., details of your tutorials, and the date of the first day of the semester is available from the study centre. You need to gather together all this information in one place, such as your diary or a wall calendar. Whatever method you choose to use, you should decide on and write in your own dates for working each unit.

Once you have created your own study schedule, do everything you can to stick to it. The major reason that students fail is that they get behind in their course work. If you get into difficulties with your schedule, please let your tutor know before it is too late for help.

Turn to Unit 1 and read the introduction and the objectives for the unit. Assemble the study materials. Information about what you need for a unit is given in the `introduction' at the beginning of each unit. You will also need both the study unit you are working on and one of your set books on your desk at the same time.

Work through the unit. The content of the unit itself has been arranged to provide a sequence for you to follow. As you work through the unit, you will be instructed to read sections from your set books or other articles. Use the unit to guide your reading.

Up-to-date course information will be continuously delivered to you at the study centre.

Work before the relevant due date (about 4 weeks before due dates), get the Assignment File for the next required assignment. Keep in mind that you will learn a lot by doing the assignments carefully. They have been designed to help you meet the objectives of the course and, therefore, will help you pass the examination. Submit all assignments no later than the due date.

Review the objectives for each study unit to confirm that you have achieved them. If you feel unsure about any of the objectives, review the study material or consult your tutor.

When you are confident that you have achieved a unit's objectives, you can then start on the next unit. Proceed unit by unit through the course and try to pace your study so that you keep yourself on schedule.

When you have submitted an assignment to your tutor for marking do not wait for its return before starting work on the next unit. Keep to your schedule. When the assignment is returned, pay particular attention to your tutor's comments, both on the Tutor-Marked Assignment form and also the written assignment. Consult your tutor as soon as possible if you have any questions or problems.

After completing the last unit, review the course and prepare yourself for the final examination. Check that you have achieved the unit objectives (listed at the beginning of each unit) and the course objectives (listed in this Course Guide).

## Tutors and Tutorials

There are some hours of tutorials (2-hour sessions) provided in support of this course. You will be notified of the dates, times and location of these tutorials. Together with the name and phone number of your tutor, as soon as you are allocated a tutorial group.

Your tutor will mark and comment on your assignments, keep a close watch on your progress and on any difficulties, you might encounter, and provide assistance to you during the course. You should mail your Tutor-Marked Assignments to your tutor well before the due date (at least two working days are required). They will be marked by your tutor and returned to you as soon as possible.

Do not hesitate to contact your tutor by telephone, e-mail, or discussion board if you need help. The following might be circumstances in which you would find help necessary.

Contact your tutor if you:

- do not understand any part of the study units or the assigned readings
- have difficulty with the self-assessment exercises
- have a question or problem with an assignment, with your tutor's comments on an assignment or with the grading of an assignment.

You should try your best to attend the tutorials. This is the only chance to have face to face contact with your tutor and to ask questions which are answered instantly. You can raise any problem encountered in the course of your study. To gain the maximum benefit from course tutorials, prepare a question list before attending them. You will learn a lot from participating in discussions actively.

## Summary

On successful completion of the course, you would have developed critical thinking skills with the material necessary for efficient and effective discussion of economic issues and integration of past events with the present. However, to gain a lot from the course please try to apply anything you learn in the course to term papers writing in other economic development courses. We wish you success with the course and hope that you will find it both interesting and useful.

**MAIN COURSE**

## CONTENTS PAGE

# Module 1: Introductory Microeconomics

Unit 1 Introduction to Microeconomics
Unit 2 Comparative Static and Dynamic Analysis

## Unit 1: Introduction to Microeconomics

**Contents**



### 1.1 Introduction

Economics is divided into two main branches: microeconomics and macroeconomics. Microeconomics deals with the behaviour of individual economic units. These units include consumers, workers, and investors, owners of land and business firms. Microeconomics explains how and why these units make economic decisions e.g. it explains how consumers make purchasing decisions and how their choices are affected by changing prices and incomes. It also explains how firms decide how many workers to hire and how workers decide where to work, how much work to do and how much to be paid for work done. Another important concern of microeconomics is how economic units interact to form larger units.

Macroeconomics deals with aggregate economic quantities, such as the level and growth rate of national output, interest rate, unemployment and inflation. Much of macroeconomics is about limits that is, the limited incomes that consumers can spend on goods and services, the limited budgets and technical know-how that firms can use to produce things, and the limited number of hours in a week the workers can allocate to labour or leisure.

## 1.2 Objectives

At the end of this unit, you should be able to:

- differentiate between microeconomics and macroeconomics
- differentiate between positive and normative economics
- explain the difference between nominal and real prices
- have broad understanding of the market concept.

## 1.3 Main Content

### 1.3.1 Positive versus Normative Economics

Microeconomics is concerned with both positive and normative questions. Positive questions deal with explanation and prediction, while normative questions deal with "*what ought to be*". In a situation where the Nigerian government imposes a quota on the importation of foreign cars; what will happen to the price, production and sales of cars? What impact will this policy change have on Nigerian consumers and on workers in the automobile industry? These questions belong to the realm of positive analysis; statements that describes relationships of cause and effect.

Positive analysis is central to microeconomics, as theories are developed to explain phenomena, tested against observations, and used to construct models from which predictions are made, all of which are vital to firms' prediction ability. For example, assuming the federal government is considering raising the tax on petroleum. The change would affect the price of petroleum, consumers' purchasing choices for small or large cars, the amounts of driving that people do, and so on. To plan sensibly, oil companies, automobile companies, producers of automobile parts, and firms in the tourism industry would all need to estimate the impact of the change. Government policy makers will also need quantitative estimate of the effects. They would want to know the cost implications on consumers, the effects on profits, and employment in the oil companies, automobile, and the tourism industries; and the amount of tax revenue likely to be collected each year.

However, sometimes we want to go beyond explanation and prediction to ask such important questions as "*what is best*". This involves **normative analysis**, which is also important for both managers of firms and those making public policy. Similarly, consider a new tax on petroleum. Automobile companies would want to determine the best (profit-maximising) mix of large and small cars to produce once the tax is in place. Specifically, how much money should be invested to make cars more fuel-efficient? For policy makers, the primary issue is likely to be whether the tax is in the public interest. The same policy objectives (say, an increase in tax revenues and a decrease in dependence on imported oil) might be met more cheaply with a different kind of tax, such as tariff on imported oil.

**Self-Assessment Exercise**

Define and differentiate between positive and normative economics.

### 1.3.2 What is a Market?

Individual economic units can be divided into two according to their functions; buyers and sellers. Buyers include the consumers, who purchase goods and services, and the firm, which buys labour, capital, and raw materials that they use to produce goods and services. Sellers include the firms, which sell their goods and services; workers, who sell their labour services; and resources owners, who rent land or sell mineral resources to firms.

The combination of buyers and sellers forms the market. A market is the collection of buyers and sellers that, through their actual or potential interactions, determine the price of a product or set of products. Also, the term market refers to economic institutions for the commercialisation of goods and services. It does not basically imply a location where goods and services are physically exchanged. It implies the summation of all transactions involving a specific good or service.

In the market for personal computers, for example, the buyers and business firms, households and students; the sellers are Hewlett-Packard (HP), Lenovo, Dell, Apple, and a number of other firms. Note that a market includes more than an industry. An industry is a collection of firms that sell the same or closely related products. In effect, an industry is the supply side of the market.

The motivating force that makes the market mechanism to work is the conflicting interests of selfish buyers and sellers and competition. Conflicts stems from the fact that the two basic sides of the market have conflicting interests. Suppliers want to sell at the highest possible prices, driven by the profit motive. Buyers on the other hand, want to purchase at the lowest possible price, driven by the utility motive. In a market, all participants want to buy cheap and sell dear. As the sellers push for the highest possible price and buyers lobby for the lowest; the resulting conflict between the basic sides of the market is resolved by competition. In general, prices tend to rise when goods become scarce, and when they are in abundance, prices tend to fall because competition will beat the price down. The twin forces of competition and self-interest make the market to operate effectively.

Sellers will always strive to obtain the highest possible prices for their products, bearing in mind that if the price is too high, competitors will make all the sales. Buyers lobby for the lowest possible price while noting that goods will in the end go to the highest bidder. The prices in most markets are determined in this manner by these market forces.

### a. Supply, Demand and Market Equilibrium

All markets are characterized by market participants making decisions to improve their own self-interest. Buyers are assumed to be motivated primarily by the desire to increase their personal satisfaction in life. Sellers are assumed to be motivated by the desire to make the largest profits possible. This interaction between buyers and sellers results in equilibrium in the market.

Equilibrium is defined to be the point at which:

$Q_s = Q_d$

Where $Q_s$ = quantity supplied, $Q_d$ = quantity demanded.

If the market price is not at equilibrium, market forces drive the market toward equilibrium. If the market price is at equilibrium, it is assumed that there exists no market pressure or incentives to move to some other level. To better understand this concept, consider the Figure 1.1.

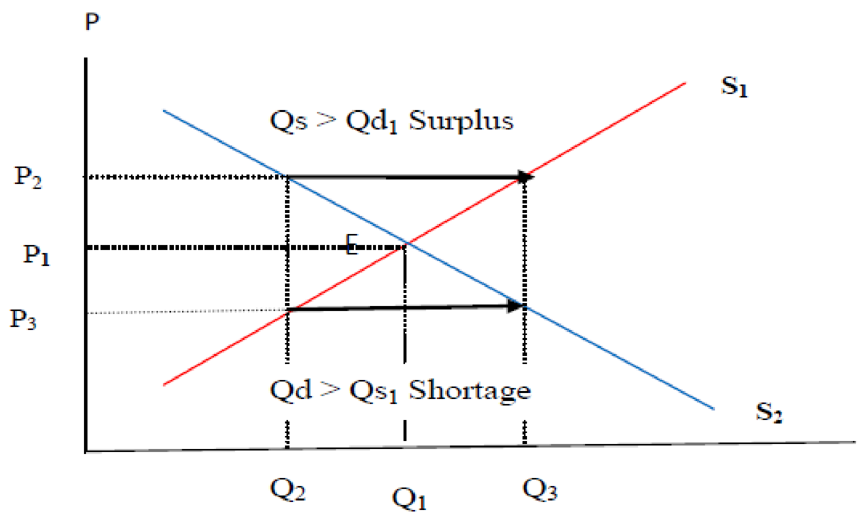



**Fig. 1.1: Supply-Demand (Market) Equilibrium**

The market represented in Figure 1.1 is in equilibrium at point E because $Q_s = Q_d$ at the common price of $P_1$. To better understand the concept of equilibrium, it is useful to examine the market at other prices. First, consider a market price above the equilibrium value. At $P_2$, the quantity supplied, $Q_3$, is larger than the quantity demanded, $Q_2$. At this price, sellers are willing to provide more to the market than consumers are willing to buy. This situation is defined as a surplus, this implies that supply exceeds demand. To combat their rising inventories (unsold stock), sellers begin to lower their prices, attempting to entice consumers to buy more. In addition, because of the lower price, sellers decrease the amount they offer for sale. As this process unfolds, there is an increase in $Q_d$ and a decrease in $Q_s$ as the market moves to the equilibrium at point E.

Now consider a market price below the equilibrium value. At the price $P_3$, the quantity demanded, $Q_3$, exceeds the quantity supplied, $Q_2$, resulting in a shortage in the market. In response to the shortage, consumers begin to beat up the price. To understand this process, consider the process of an auction. An auction usually begins with the number of willing and able buyers exceeding the amount of product available. The role of the auctioneer is to bid the price up until $Q_s = Q_d$. As the price increases, consumers begin to drop out of the bidding process. In the market represented above, consumers begin bidding the price up and, as a result, some consumers begin to drop out of the market and $Q_d$ begins to fall. As the price begins to increase, sellers who are willing and able respond by increasing the amount they provide to the market, and $Q_s$ begins to increase. These changes in $Q_s$ and $Q_d$ continue until the market reaches equilibrium at a price of $P_1$ and a quantity of $Q_1$.

Ultimately, markets tend to move toward equilibrium. When market prices are too high, surpluses exist, and market forces drive the prices down to the equilibrium level. When market prices are too low, shortages exist, and market forces drive up the prices to the equilibrium price. At equilibrium, there exists no market pressure for change, and thus the price tends to stay at the equilibrium until something in the economy changes, resulting in a change in either supply or demand.

However, there exists a debate as to whether market equilibrium exists; and if it does, what is its relevance to the choices made by firms and consumers in the market place? Some have argued that the essential build up to this debate is based on the doubt created by the fact that the unique price at equilibrium presupposes that there is a seemingly obvious possibility that some potential consumers might be denied access to the economy’s output because their income could be less than the market price. The relevance of the equilibrium price can be understood from how it underlays the characterization of the choice of the rate and level of employment of human and non-human resources by firms and expenditure switching by consumers in allocating income. Also, the pursuit of the profit motive by firms cast doubt on the assumption that at equilibrium there exist no incentive for any change in behavior by consumers and businesses.

**b. Quantitative Approach to Market Equilibrium**

The law of demand suggests that an inverse mathematical relationship exists between price and quantity demanded. The law of supply suggests that a similar but positive relationship exists between price and quantity supplied. An example of these relationships is expressed below.

(i) The case of a single product (X)

Given a simple linear demand and supply function

$Qd= 6 – 0.5p$ (1)

$Qs= -4 + 1.5p$ (2)

At equilibrium, $Qs = Qd \; => Qe= Qd – Qs$

Equating (1) and (2),

$6 – 0.5p =-4 + 1.5p \; p=5$

Substituting value of p into (1) or (2)

$Qe= 3.5$

Thus, equilibrium price and quantity are Pe = 5 and Qe = 3.5 (ii)

The case of two related goods (x and y)

Given two substitute goods (X and Y) with Demand and Supply functions,

$Qd_x= 41 - 1.5p_x + 0.5p_x$ (3)

$Qs_x = -2.5 + 7.5p_x$ (4)

$Qd_y = 46 + p_x – 2p_y$ (5)

$Qs_y= -3 + 16p_y$ (6)

At equilibrium,

$Qd_x - Qs_x = Qe_x$

$Qd_y - Qs_y = Qe_y$

Equating (3) and (4)

$41 - 1.5p_x + 0.5p_y = -2.5 + 7.5p_x$

$9p_x - 0.5p_y = 43.5 \qquad (7)$

Equating (5) and (6)

$46 + p_x - 2p_y = -3 + 16p_y$

$-p_x + 18p_y = 49 \qquad (8)$

Equating (7) and (8) gives us a simultaneous equation system. Solving simultaneously,

$P_x = 5$ *and* $p_y = 3$

Substituting value of $p_x$ into (3) or (4),

$Qe_x = 35$

Substituting value of $p_y$ into (5) or (6)

$Qe_y = 45$

Thus, equilibrium price and quantity for each product is

$P_x = 5, Qe_x = 35$

$P_y = 3, Qe_y = 45.$

### c. Market Failure

Microeconomic theory is founded on the premise of a smooth functioning market. The market (price) system allocates resources by providing information to resources owners on the relative price of individual goods and services. The price mechanism is a system whereby the information provided by the relative prices of goods and services determines freely what the society is to produce, how and for whom it is to be produced. The price system is assumed to allocate resources effectively, bringing maximum benefits to the society and resources owners.

There are three main arguments proffered in support of the price system as a means of resource allocation. They are:

(i) Its ability to achieve optimality in the allocation of societal resources.
(ii) Its built-in flexibility which negates the need for conscious government planning and intervention.
(iii) Its guarantee of economic freedom to all and sundry.

Market failure is a phrase that is usually employed to describe one or both of two things:

(i) The failure of the price mechanism, at any point in time, to achieve optimal resource allocation.
(ii) The inability of the price system to achieve other desirable societal goals associated with the distribution and utilization of resources.

In relation to optimal resource allocation and utilization, the price mechanism does not, by itself, accommodate externalities in economic transactions. Finally, the price mechanism breaks down whenever and wherever there are inefficient transportation and communication systems. With market imperfections, there is bound to be resource unemployment and imbalances within the economy.

**Self-Assessment Exercise**

Explain the concept of market, market equilibrium and market failure.

### 1.3.3 Competitive versus Non-Competitive Markets

#### a. Competitive Market

A perfectly competitive market has many buyers and sellers, so that no single buyer or seller has a significant impact on price. Most agricultural markets are close to being perfectly competitive. For example, thousands of farmers produce maize, which thousands of buyers purchase to produce food and other products. As a result, no single farmer and no single buyer can significantly affect the price of maize. Equilibrium occurs in a competitive market when MR = MC = P (MR = marginal revenue, MC = marginal cost, P = price), which can lead to Pareto's optimum. Firms in a competitive market are price takers because they produce such a small share of their industries output, and their production decisions have little impact on the total quantity of output sold.

**b. Marginal Revenue (MR) and Marginal Cost (MC)**

Marginal revenue indicates how much extra revenue is received for selling an extra unit of output. Marginal revenue plays a key role in the profit maximizing decision of a perfectly competitive firm relative to marginal cost (MC). A perfectly competitive firm maximizes profit by equating marginal revenue, the extra revenue generated from production, with marginal cost (MC) *which is the cost incurred in the production of extra output.* If these two marginals are not equal, then profit can be increased by producing more or less output. In other words, when MR is greater than MC, the decision is for the firm to increase output; and when the MR is less than the MC, the firm should reduce output. Thus, the firm maximizes profit at that output point where MR = MC.

**c. Non-Competitive Market**

What if the industry has only one firm? How does the firm behave when it is a monopolist that differs from its behaviour in a competitive industry? Because the industry has only one firm, that firm must consider the effect of its production decisions on the price of its output. The problem is that when the monopolist produces an additional unit of output it **cannot** sell that output at the prevailing price because the prevailing price corresponds to some point on the demand curve for the product. At that price, the most consumers are willing to buy some amount, say $q_0$. To sell an addition unit of output, the monopolist must cut its price by an amount P for all $q_0$units of output. This behaviour implies that the marginal revenue associated with selling an additional unit of output is less than the price that this unit of output is sold for in the market:

$MR = P \ \ + P *q_0 < P$, because $P < 0$

Intuitively, the monopolist will maximize profits by raising prices. But to do so, it must reduce output. This is the same problem that OPEC faces when trying to control oil prices. In order for a $90 per barrel price to prevail on world markets each producer must agree to cut back production. The reason that OPEC is relatively ineffective is because it cannot discipline itself to keep production within preset limits.

**Self-Assessment Exercise**

Define and differentiate between competitive and non-competitive market, and give an example of both.

### 1.3.4 Real versus Nominal Prices

We often want to compare the price of goods today with what it was in the past or is likely to be in the future. To make such a comparison meaningful, we need to measure prices relative to an overall price level. In absolute terms, the price of a dozen of oranges is many times higher today than it was 50 years ago. We need to correct for inflation when comparing prices across time. This means measuring prices in real rather than nominal terms. For example, the nominal value of a pen is its money values in different years. Real values adjust for differences in the price level in those years i.e. Gross Domestic Product (GDP) and income. For a series of nominal values in successive years, different values could be because of differences in general price level. Nominal value does not specify how much of the differences are from changes in price level. Real values remove this ambiguity.

Real values convert the nominal values as if prices were held constant in each year of the series. Any differences in real values are then attributed to differences in quantities of the good that the money income could buy in each year. For example, suppose the inflation rate in Nigeria is 3% in 2022. We can buy a basket of onion today (2022) and it will cost 100 naira, or we can buy that basket next year (2023) and it will cost 103 naira. If we buy a bond with a 6% nominal interest rate for 100 naira in 2022, sell it after a year (2022) and get 106 naira, buy a basket of onion for

103 naira in 2023 from the proceed of the bond, we will have 3 naira left over. So, after factoring in inflation, our 100 Naira bond will earn us 3 naira in income in 2023.

**a. Nominal and Real Values**

The nominal value of a commodity bundle in a given year depends on both quantities and then current prices.

Nominal value/ Real value = $\frac{(P*Q)}{Q} = P$

Here, P is a price index, and Q is a quantity index of real value. P is constructed to be equal to 100 in the base year. Alternatively, P can be constructed to equal 100 in the base year.

Consider a simple case of one commodity. Here, output and consumption may be measured either in terms of money value (nominal) or physical quantity (real). Let us denote that commodity and let:

P*i* = the unit price of *i*, say ₦5

Q*i* = the quantity of good *i*, say 10 units.

The nominal value of the good would then be price * quantity. Nominal value of good i = Pi * Qi = ₦5 * 10 = ₦50.

Given only the nominal value and price, derivation of a real value is immediate.

Real value of good i = $\frac{(Pi*Qi)}{Pi} = Qi$ = 50/5 = 10.

The price "deflates" (divide) the nominal value to derive a real value.

**b. Nominal and Real Price**

Economic decisions are mostly based on relative prices, not absolute prices. In other words, if all prices and income doubled, there would be no effect on purchasing decisions because relative prices would not have changed. Similarly, planting decisions of farmers are largely determined by the relative prices of different crops.

Inflation describes a general increase in all prices, although the rate of increase varies across products. Inflation is usually measured by the consumer price index (CPI), which describes the prices in a given month as a percentage of prices in a base period. If the CPI has a base month of January 2000, then by definition the CPI for that month is 100. If the CPI is 150 in January 2002, this means that prices were 50 per cent higher in January 2002 than they were two years before. This implies that the annual inflation rate over this period is (1.5)1/2 = 1.225 or 22.5%.

Real price in month t $=\frac{Pt}{\left(\frac{CPIt}{100}\right)}$

**Real prices** are defined as prices that have been adjusted for inflation. The real price in a given month is calculated by dividing the nominal price (the price observed in the market) by the CPI of that month, where the CPI is expressed as a ratio and not a percentage. In other words, a CPI of 150 is expressed as 1.5. Usually, we are most interested in the changes in real prices over time rather than the level of real prices at one time. For example, it is not very useful to know that the real price of maize in June 2008 was say ₦40 per unit. It is more interesting to know that, although nominal prices of maize in 2008 increased by 80%, the real prices of maize increased by only 45%. This implies that, relative to other prices in the economy, the price of maize rose by 45% over the year.

**Self-Assessment Exercise**

In what way is nominal price different from real price?

## 1.4 Conclusion

In this unit, we differentiated between microeconomics and macroeconomics. Microeconomics focuses on individual unit while macroeconomics deals with aggregates. Positive and Normative economics both answers the question of “what is” and “what ought to be”, as positive economics deals with explanation and predictions, while normative economics answer the question of what ought to be and provision of alternative options. Market comprises of both the buyers and sellers, and through their interactions, prices of products are determined. When the quantity demanded by the consumer is equal to the quantity supplied by the seller, the market is in equilibrium. We concluded that market failure is likely to occur when the price system is unable to achieve optimal resource allocation through information asymmetry.

## 1.5 Summary

This unit treats the fundamentals of microeconomics by looking into the concept of positive and normative economics as well as markets. Market equilibrium was analyzed using both algebra and graphical presentations and finally, market failure. We realised that it is necessary for students to be refreshed on these basic topics and terms before moving ahead to other units and modules in order to facilitate easy understanding of the concepts and terms used.

## 1.6 Tutor-Marked Assignment

1. Calculate the real price of an apple in 1980 naira.
   a. Has the real price increased, decreased or stayed the same from 1980 to 2010?
2. Explain the concept positive and normative economics.
3. Explain the concept of real and nominal price.
4. a. What do you understand by market equilibrium?
   b. What do you understand by market failure?
5. a. What is competitive and non-competitive market?
   b. Give examples of each.
6. Analyse the competitive market equilibrium using figures.
7. Given the demand and supply functions $Qd_x = 25 - 10_p$ and $Qs_x = -5 + 3_p$ Determine the equilibrium price and quantity.

## 1.7 References/Further Reading

Krugman, P. & Robin, W. (2012). *Economics*. (3rd ed.). New York: Worth Publishers.

Lipsey, R. et al. (2008). *Economics*. (13th ed.). New Jersey: Pearson.

Mankiw, N. G. (2009). *Principles of Economics*. (7th ed.). Mason, OH: South-Western Cengage Learning.

Robert, S. P. & Daniel, L. R. (2009). *Microeconomics*. (7th ed.). Pearson Education International.

Sloman, J. (2006). *Economics.* (6th ed.). Prentice Hall

## 1.8 Answers to Self-Assessment Exercises

**Question**

Define and Differentiate between positive and normative economics.

**Model Answer**

Positive economics deal with explanation and prediction, while normative economics deal with "*what ought to be*". Positive economic analysis is central to microeconomics, as theories are developed to explain phenomena, tested against observations, and used to construct models from which predictions are made. However, sometimes we want to go beyond explanation and prediction to ask such important questions as "*what is best*". This involves normative analysis, which is important for firms and public policy.

**Question**

Explain the concept of market, market equilibrium and market failure.

**Model Answer**

Individual economic units can be divided into two according to their functions; buyers and sellers. Buyers include the consumers, who purchase goods and services, and the firm, which buys labour, capital, and raw materials that they use to produce goods and services. Sellers include the firms, which sell their goods and services; workers, who sell their labour services; and resources owners, who rent land or sell mineral resources to firms. The combination of buyers and sellers forms the market. A market is the collection of buyers and sellers that, through their actual or potential interactions, determine the price of a product or set of products. Also, the term market refers to economic institutions for the commercialisation of goods and services. It does not basically imply a location where goods and services are physically exchanged. It implies the summation of all transactions involving a specific good or service.

Market equilibrium is defined as that point where quantity supplied of a good or is equal to the quantity demanded of that good. If the market price is not at equilibrium, market forces drive the market toward equilibrium. If the market price is at equilibrium, it is assumed that there exists no market pressure or incentives to move to some other level. Market failure is a phrase that is usually employed to describe one or both of two things: First, the failure of the price mechanism, at any point in time, to achieve optimal resource allocation. Second, the inability of the price system to achieve other desirable societal goals associated with the distribution and utilization of resources.

**Question**

Define and differentiate between competitive and non-competitive market, and give an example of both.

**Model Answer**

A competitive market has many buyers and sellers, so that no single buyer or seller has a significant impact on price. Most agricultural markets are close to being perfectly competitive. For example, thousands of farmers produce maize, which thousands of buyers purchase to produce food and other products. As a result, no single farmer and no single buyer can significantly affect the price of maize.

Non-competitive market has a single seller and many buyers. Consequently, the seller has a significant influence on the price of its product. This is monopoly power, a case in which the firm constitutes the industry. Because the industry has only one firm, that firm must consider the effect of its production decisions on the price of its output. The problem is that when the monopolist produces an additional unit of output it cannot sell that output at the prevailing price because the prevailing price corresponds to some point on the demand curve for the product.

**Question**

In what way is nominal price different from real price?

**Model Answer**

Real price is defined as price that has been adjusted for inflation. The real price in a given month is calculated by dividing the nominal price (the price observed in the market) by the CPI of that month, where the CPI is expressed as a ratio and not a percentage.

# Unit 2: Comparative Static and Dynamic Analysis

**Contents**



## 2.1 Introduction

Comparative static and dynamic analyses are commonly used to study changes in supply and demand when analyzing a single market, and to study changes in monetary or fiscal policy when analyzing the whole economy. Comparative static simply compares different equilibrium state, while dynamic analysis studies the market disequilibrium state and its adjustment process.

## 2.2 Objectives

At the end of this unit, you should be able to:

- explain and differentiate between static and dynamic analysis
- analyze the Cobweb model
- analyze the static and dynamic Cobweb diagrams.

## 2.3 Main Contents

### 2.3.1 Comparative Static

We are often interested in analyzing how the consumer's choice varies with changes in his wealth and in prices. The examination of a change in outcome in response to a change in underlying economic parameters is known as comparative static analysis.

Comparative static analysis is the comparison of two different equilibrium states, after the process of adjustment. It does not study the motion towards equilibrium, nor the process of the change itself.

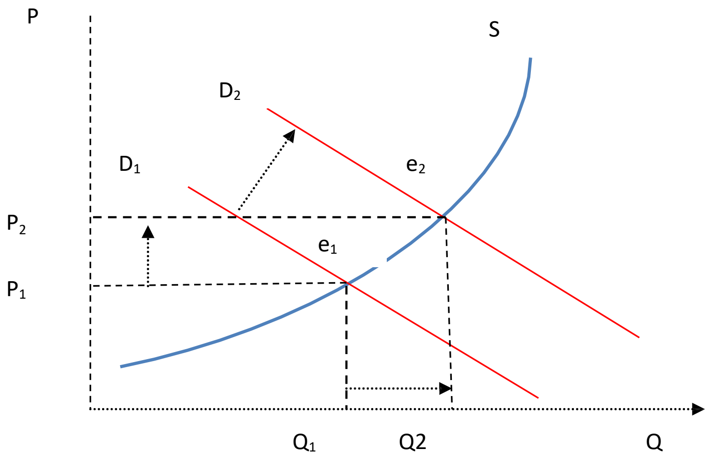



**Fig. 1.2: Comparative Static Equilibrium**

In Figure 1.2, comparative statics shows an increase in demand causing a shift in price and quantity. In comparing two equilibrium states, comparative statics does not describe how the increases actually occur.

Using Figure 1.2 for example, the initial equilibrium position, price and quantity for the market is $e_1$, $P_1$ and $Q_1$ respectively. Assuming there is an expansion in demand over time, this will lead to a new equilibrium at $e_2$, which is higher than the initial equilibrium. Under comparative statics, the focus is on comparing these two equilibriums without regard to the process of adjustment from one equilibrium to another.

### a. Linear Approximation

Comparative statics results are usually derived by using the implicit function theorem to calculate a linear approximation to the system of equations that defines the equilibrium, under the assumption that the equilibrium is stable. That is, if we consider a sufficiently small change in some exogenous parameter, we can calculate how each endogenous variable changes using only the first derivatives of the terms that appear in the equilibrium equations.

For example, suppose the equilibrium value of some endogenous variable *x* is determined by the following equation:

$$f(x, a) = 0$$

Where α is an exogenous parameter. Then, given a first-order approximation, the change in x caused by a small change in *α* must satisfy:

$$B\mathrm{d}x + C\mathrm{d}a = 0$$

Here $\mathrm{d}x$ and $\mathrm{d}a$ represent the changes in $x$ and $a$, respectively, while $B$ and $C$ are the partial derivatives of $f$ with respect to $x$ and $a$ (evaluated at the initial values of $x$ and $a$), respectively.

Equivalently, we can write the change in $x$ as:

$$\mathrm{d}x = -B^{-1}C\mathrm{d}a.$$

The elements of $-B^{-1}C$ are sometimes called the multipliers of the elements of *a* on the elements of *x*.

**b. Application for Profit Maximization**

Suppose a firm produces *n* goods in quantities $x_1, \dots, x_n$. The firm's profit is a function *p* of $x_1, \dots, x_n$ and of *m* exogenous parameters $q_1, \dots, q_m$ which may represent, for instance, various tax rates. Provided the profit function satisfies the smoothness and concavity requirements; the comparative statics method above describes the changes in the firm's profit due to small changes in the tax rates.

**c. Limitations of Comparative Statics**

One limitation of comparative statics using the implicit function theorem is that results are valid only in a (potentially very small) neighborhood of the optimum-that is, only for very small changes in the exogenous variables. Another limitation is the potentially overly restrictive nature of the assumptions conventionally used to justify comparative statics procedures.

**Self-Assessment Exercise**
Explain the comparative statics model and its application for profit maximization.

### 2.3.2 Dynamic Analysis

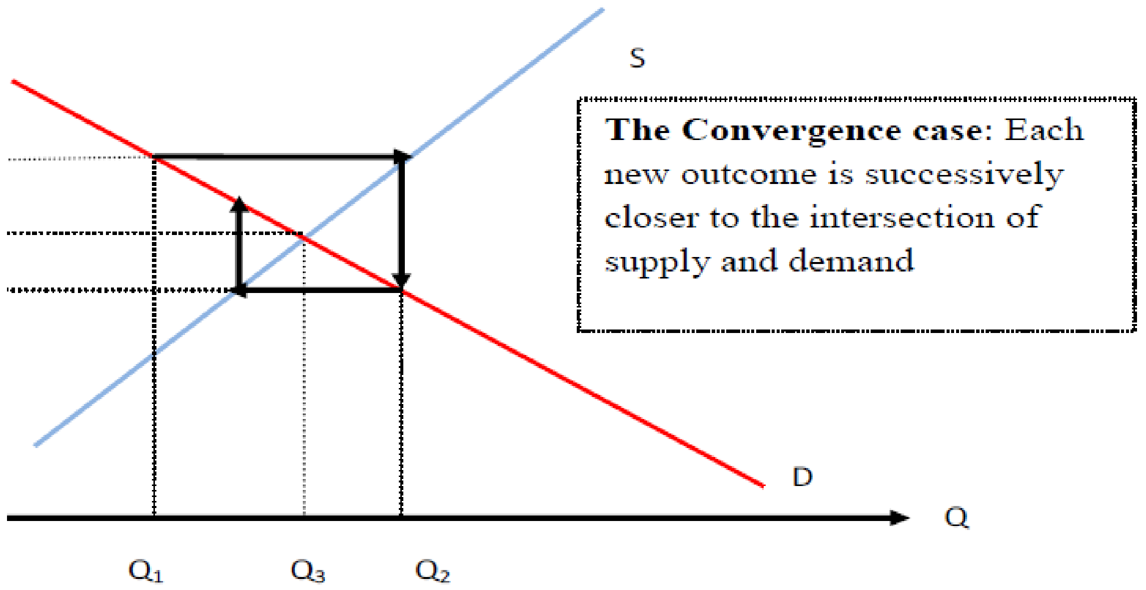
Supply often does not respond freely to price changes. If there is an increase in the price of a product, except where there is a stock of inventory, producers must first organize the factors of production towards producing more of the commodity before there can be an increase in supply. Thus, there is a time lag in supply's response to price changes. Dynamic analysis studies the time part of market disequilibrium or fluctuation and the possibility of market adjustment process towards equilibrium over time.

A simple dynamic model to be considered is a situation where supply responds to price with a time lag (one period of time), while demand responds to price in the same period without a time lag nor in anticipation of future price. Such a model is known as the Cobweb model. It reflects certain markets in reality such as the agricultural produce or the crop market where supply cannot be changed at short notice.

**The Cobweb Model**

The cobweb model or cobweb theory is an economic model that explains why prices might be subject to periodic fluctuations in certain types of markets. It describes cyclical supply and demand in a market where the amount produced must be chosen before prices are observed. Producers' expectations about prices are assumed to be based on observations of previous prices. The model shows how achieving supply and demand equilibrium might be so automatic if, as seems reasonable, the suppliers set the price and the consumers react with a quantity demanded. For some of the demand and supply curves, the equilibrium can be unstable.

The cobweb model is the classic demonstration that dynamic behaviour by economic agents might not converge to a stable equilibrium with supply equal to demand. This application provides two ways to show the outcome. We can experiment with the key parameter that determines whether the outcome is stable or not.



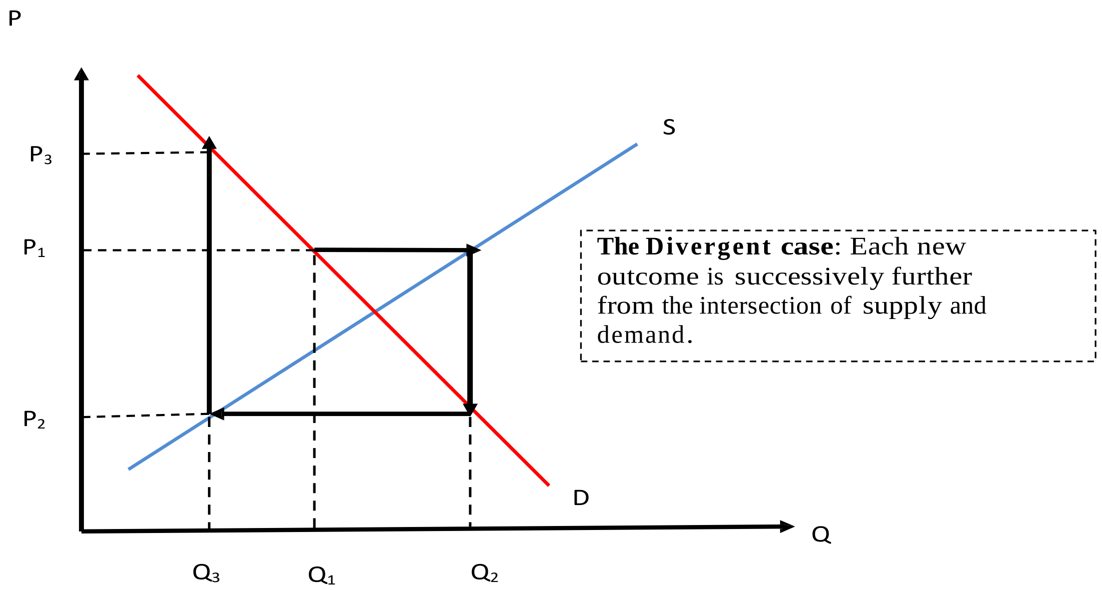



**Fig. 1.3: The Cobweb Model**

## The Model

The cobweb model is based on a time lag between supply and demand decisions. Agricultural markets are a context where the cobweb model might apply, since there is a lag between planting and harvesting. Suppose that as a result of unexpectedly bad weather, farmers go to market with an unusually small crop of produce. This shortage, equivalent to a leftward shift in the market's supply curve, results in high prices. If farmers expect these high price conditions to continue, then in the following year, they will raise their production relative to other crops. Therefore, when they go to market the supply will be high, resulting in low prices. If they then expect low prices to continue, they will decrease their production for the next year, resulting in high prices again. This process is illustrated in the first diagram in figure 1.3. The equilibrium price is at the intersection of the supply and demand curves. A poor harvest in Period 1 means supply falls to $Q_1$, so that prices rise to $P_1$. If producers plan their Period 2 production under the expectation that this high price will continue, then Period 2 supply will be higher, at $Q_2$. Prices therefore fall to $P_2$ when they try to sell all their output. This process repeats itself, oscillating between periods of low supply with high prices and then high supply with low prices, leading to the price and quantity spiraling inward or outward. When price and quantity spiral inwards, as in the first diagram in figure 1.3, in which case the economy converges to equilibrium where supply and demand crosses; or spiral outwards, with the fluctuations increasing in magnitude.

Simplifying, the cobweb model can have two main types of outcomes:

(i) If the supply curve is steeper than the demand curve, then the fluctuations decrease in magnitude with each cycle, so a plot of the prices and quantities over time would look like an inward spiral, as shown in the first diagram. This is called the stable or convergent case.

(ii) If the slope of the supply curve is less than the absolute value of the slope of the demand curve, then the fluctuations increase in magnitude with each cycle, so that prices and quantities spiral outwards. This is called the unstable or divergent case.

Two other possibilities are:

(i) Fluctuations may also remain constant in magnitude, so a plot of the outcomes would produce a simple rectangle, if the supply and demand curves have exactly the same slope (in absolute value).

(ii) If the supply curve is less steep than the demand curve near the point where the two curves cross, but steeper when we move sufficiently far away, then prices and quantities will spiral away from the equilibrium price but will not diverge indefinitely; instead, they may converge to a limit cycle.

**Self-Assessment Exercise**

Define dynamic analysis and explain the cobweb model.

## 2.4 Conclusion

In this unit, you have been introduced to the concepts of comparative static and dynamic analysis. You also understand the difference between static and dynamic analysis. Linear approximation was calculated using implicit function before its application for profit maximisation was treated. It is believed that the cobweb model analysis of equilibrium and the limitations of comparative statics are well understood.

## 2.5 Summary

The concepts of comparative static and dynamic analysis were discussed. A diagram was presented to analyse the comparative static model as well as the convergence and divergence diagrams of the cobweb model which explains the stability and instability of markets.

## 2.6 Tutor-Marked Assignment

1. What is comparative static analyses?
2. What is dynamic analysis?
3. Differentiate between static and dynamic analysis
4. Draw and analyse the comparative static and dynamic cobweb diagram.

## 2.7 References/Further Reading

Frank, R. H. & Bernanke, B. (2013). *Principles of Economics*. (5th ed.). New York: McGraw-Hill.

Sexton, R. L. (2013). *Exploring Economics*. (6th ed.). Mason, OH: South-Western Cengage Learning.

Rittenberg, L. & Tregarthen, T. (2012). *Principles of Economics*. V2.0. Washington, DC: Flat World Knowledge.

Robert, S. P. & Rubinfeld, D. L. (2009). *Microeconomics.* (7th ed.). New Jersey: Pearson Education International.

## 2.8 Answers to Self-Assessment Questions

**Question**

Explain the comparative statics model and its application for profit maximization.

**Model Answer**

The examination of a change in outcome in response to a change in underlying economic parameters is known as comparative static analysis. Comparative static analysis is the comparison of two different equilibrium states, after the process of adjustment. It does not study the motion towards equilibrium, nor the process of the change itself.

The application of comparative static model for profit maximization can be illustrated as follows: suppose a firm produces *n* goods in quantities $x_1, \dots, x_n$. The firm's profit is a function *p* of $x_1, \dots, x_n$ and of *m* exogenous parameters $q_1, \dots, q_m$ which may represent, for instance, various tax rates. Provided the profit function satisfies the smoothness and concavity requirements; the comparative statics method above describes the changes in the firm's profit due to small changes in the tax rates.

**Question**

Define dynamic analysis and explain the cobweb model.

**Model Answer**

Dynamic analysis studies the time part of market disequilibrium or fluctuation and the possibility of market adjustment process towards equilibrium over time.

The cobweb model is an economic model that explains why prices might be subject to periodic fluctuations in certain types of markets. It describes cyclical supply and demand in a market where the amount produced must be chosen before prices are observed. Producers' expectations about prices are assumed to be based on observations of previous prices. The model shows how achieving supply and demand equilibrium might be so automatic if, as seems reasonable, the suppliers set the price and the consumers react with a quantity demanded.

# Module 2: Theory of Demand, Consumer Behaviour and Optimization



## Unit 1 Theory of Demand

**Contents**



### 1.1 Introduction

The purpose of the theory of demand is to determine the various factors that affect demand. Demand is a multivariate relationship, that is, it is determined by many factors simultaneously. Some of the most important determinants of the market demand for a particular product are its own price, consumer's income, prices of other goods, consumer's tastes, income, total production, consumer wealth, credit availability, government policy, past level of demand and income. Traditional theory of demand has concentrated on four of the above determinants, the price of the good, price of other goods, income and taste.

### 1.2 Objectives

At the end of this unit, you should be able to:

- identify the determinants of demand
- differentiate between individual and market demand
- differentiate between normal and inferior goods
- explain the concept of consumer surplus.

## 1.3 Main Content

### 1.3.1 Individual and Market Demand

Demand, as a basic economic principle, is the effective want for something and the willingness and ability to pay for it. As a relative concept, demand is always attached to a certain price point at a particular point in time. Quantitative demand analysis provides useful guidance to companies and investors trying to determine their market strategy and the growth potential of a product. There are two basic types of demand: individual and market. While both principles overlap in many ways, the scope of individual demand is much narrower than market demand.

#### a. Individual Demand

Individual demand is the demand of one individual or firm. It represents the quantity of a good that a single consumer would buy at a specific price point at a specific point in time. While the term is somewhat vague, individual demand can be represented by the point of view of one person, a single family, or a single household.

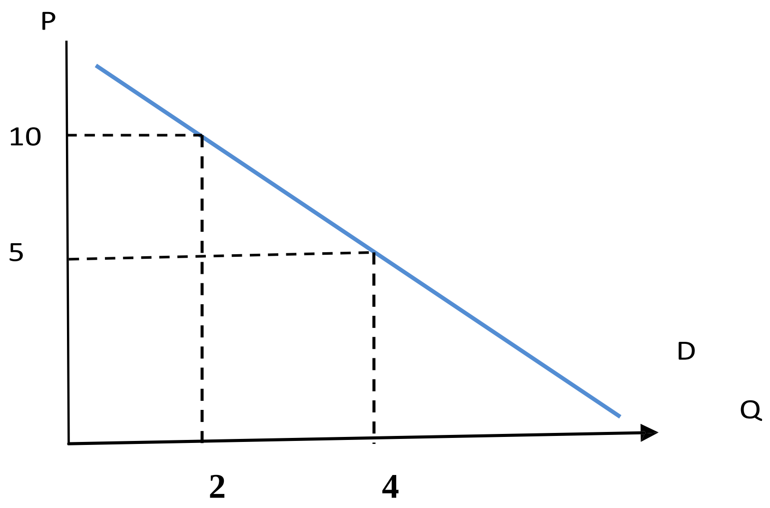


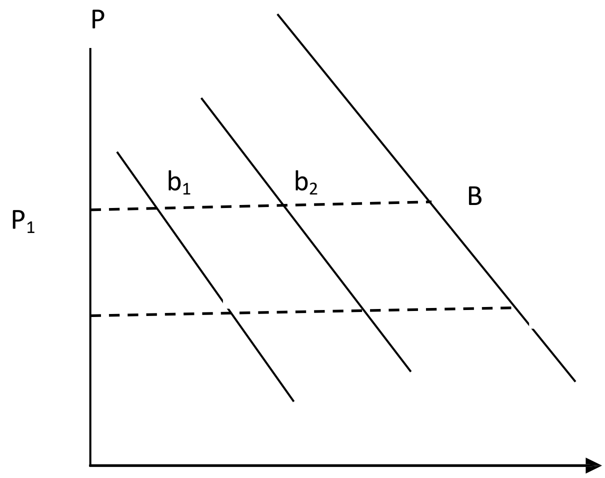
**Fig. 2.1: The Demand Curve**

Based on the above demand, it shows that an increase in price from ₦5 to ₦10 led to a fall in quantity demanded from 4 to 2 and vice versa. This depicts the individual demand curve.

**b. Market Demand**

Market demand provides the total quantity demanded by all consumers. In other words, it represents the aggregate of all individual demands. There are two basic types of market demand: primary and selective. Primary demand is the total demand for all of the brands that represent a given product or service, such as all phones or all high-end watches. Selective demand is the demand for one particular brand of product or service, such as the iPhone. Market demand is an important economic marker because it reflects the competitiveness of a marketplace, a consumer's willingness to buy certain products and the ability of a company to leverage itself in a competitive landscape. If market demand is low, it signals to a company that they should terminate a product or service, or restructure it so that it is more appealing to consumers.

In a clear manner, market demand is the combination of all individual demand. Let us assume, there are three consumers in the market who purchase the goods at various prices. The individual demand curves are $d_A$, $d_B$ and $d_C$ in Figure 2.2. When the price is $P_2$, consumer A cannot purchase any good, consumer B purchase 5 units, while consumer C bought 15 units. The total purchases at price $P_2$ are (10 + 15) = 25 units of good. If the price is P, the combined purchases of all the consumers are 35 units of good (3 + 13 + 19).



$a_2$
$a_1$
A
$P_2$
$d_2$
$D^*$
$_1d$
Q

**Fig. 2.2: Derivation of Market Demand Curve**

Figure 2.2 depicts the summation of individual demand to give market demand. The market demand was derived from the individual demand curves by horizontally summing the individual curves. At each price, we sum the quantities each consumer will buy to obtain the total demanded at their prices.

### c. Factors Influencing Demand

There are several factors that influence individual and market demand. Individual demand is influenced by an individual's age, sex, income, habits, expectations and the prices of competing goods in the marketplace. Market demand is influenced by the same factors, but on a broader scale, the taste, habits and expectations of a community and so on. It also considers the number of buyers in the market, the rate at which a certain community is growing and the level of innovation erupting in the marketplace. Market demand can be measured on an international, national, regional, local, or even smaller level.

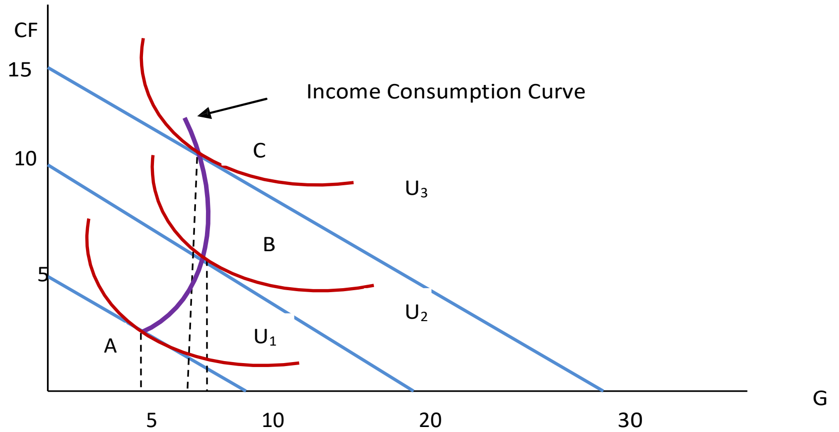
**Self-Assessment Exercise**

Define and differentiate between individual demand and market demand.

## 1.3.2 Normal versus Inferior Goods

When the income consumption curve has a positive slope, the quantity demanded increases with incomes. As a result, the income elasticity of demand is positive. The greater the shift to the right of the demand curve, the larger the income elasticity. In this case, the goods are described as *normal* (consumers want to buy more of them as their income increases).

In some cases, the quantity demanded falls as income increases; the income elasticity of demand is negative. We then describe the good as *inferior*. The term inferior simply means that consumption falls when income rises. For example, *Garri* is inferior for some people: As their income increases, they buy less *Garri* and more corn-flakes. Figure 2.3 shows the income-consumption curve for an inferior good. For relatively low income levels both *Garri* and corn-flakes are normal goods. As income rises, however, the income-consumption curve bends backward (from point B to C). This shift occurs because *Garri* has become an inferior good, and its consumption has fallen as income increased.




CF = Corn-Flakes; G = Garri

**Fig. 2.3: Income vs. Consumption**

An increase in a person's income can lead to less consumption of one of the two goods being purchased. Here, *Garri*, though a normal good between A and B, becomes an inferior good when the income- consumption curve bends backward between B and C.

**a. Indifference Curve Approach to Normal and Inferior Goods**

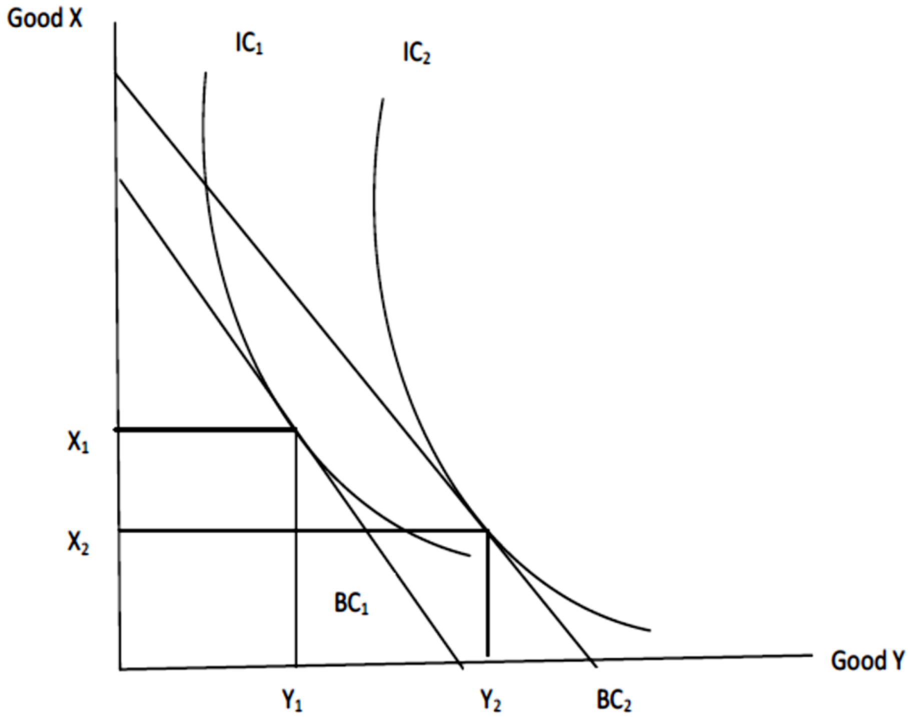



**Fig. 2.4: Indifference Curve** (IC = Indifference Curve BC = Budget Constraint)

Depending on indifference curve, the amount of a good bought can either increase, decrease, or stay the same when income increases. In Figure 2.4, good Y is a normal good since the amount purchased increases from $Y_1$ to $Y_2$ as the budget constraint shifts from $BC_1$ to the higher income $BC_2$. Good X is an inferior good since the amount bought decreases from $X_1$ to $X_2$ as income increases.

**Self-Assessment Exercise**
Describe the inferior and normal goods using indifference curve and income consumption curve

### 1.3.3 Substitutes and Complements

The demand curve shows the relationship between the price of a good and the quantity demanded, with preferences, income, and prices of other goods held constant. Examples of complementary goods are football boot and football, bread and tea, computer hardware and software, etc. other goods such as *Milo* and *Bournvita*, flour and wheat, movie ticket and DVD disk, tend to substitute for one another. Two goods are substitutes if an increase in the price of one lead to an increase in the quantity demanded of the other. If the price of a movie ticket rises, we would expect individuals to rent or buy more DVD discs, because movie ticket and DVD disc are substitutes. Similarly, two goods are complements if an increase in the price of one good leads to a decrease in the quantity demanded of another good. For example, if the price of petrol goes up, causing petrol consumption to fall, we would expect the consumption of motor oil to fall as well, because petrol and motor oil are used together. Two goods are however independent if a change in the price of one good has no effect in the quantity demanded of the other.

### 1.3.4 Consumer Surplus

Consumer surplus is the monetary gain obtained by consumers because they are able to purchase a product for a price that is less than the highest price that they would be willing to pay. The concept of consumer surplus was introduced to deal with novel complexity raised by the dominance of monopoly capital. On a supply and demand diagram, consumer surplus is the area above the equilibrium price of the good and below the demand curve. This reflects the fact that consumers would have been willing to buy a single unit of the good at a price higher than the equilibrium price, a second unit at a price below that but still above equilibrium, etc., yet they in fact pay just the equilibrium price for each unit they buy.

Put differently, consumer surplus is the difference between the maximum price a consumer is willing to pay and the actual price he does pay. If a consumer is willing to pay more than the current asking price, then they are getting more benefit from the purchased product than they spent to buy it. An example of a good with generally high consumer surplus is drinking water. People would pay very high prices for drinking water, as they need it to survive. The difference in the price that they would pay if they had to, and the amount that they pay now is their consumer surplus. Note that the utility of the first few liters of drinking water is very high (as it prevents death), so the first few liters would likely have more consumer surplus than subsequent liters.

The maximum amount a consumer would be willing to pay for a given quantity of a good is the sum of the maximum price he would be willing to pay for the first unit, the (lower) maximum price he would be willing to pay for the second unit, etc. Typically, these prices are decreasing; they are given by the individual demand curve. For a given price the consumer buys the amount for which the consumer surplus is highest,

where consumer surplus is the sum, over all units, of the excess of the maximum willingness to pay over the equilibrium (market) price. The consumer's surplus is highest at the largest number of units for which, even for the last unit, the maximum willingness to pay is not below the market price. Figure 2.5 depicts the consumer surplus.

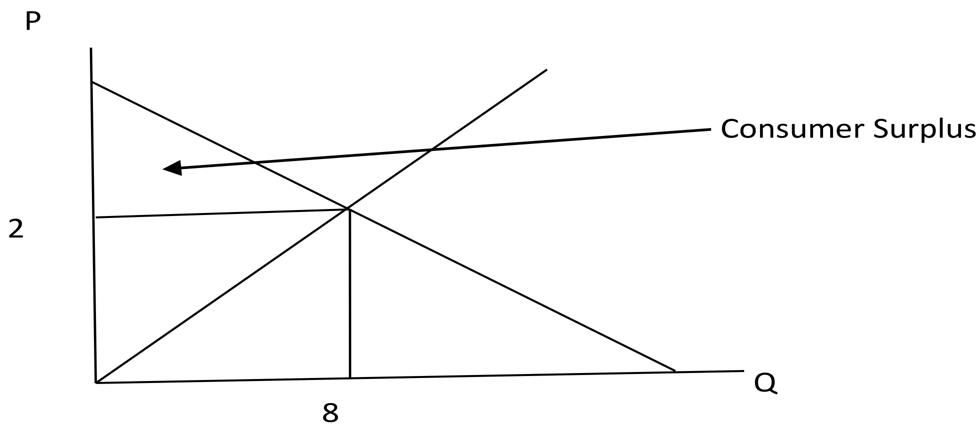



**Figure 2.5: Consumer Surplus**

**Self-Assessment Exercise**

Explain the concept of consumer surplus with graphical illustration.

## 1.4 Conclusion

From our discussion so far on the theory of demand, we can infer the following:

First, income, taste, price and other prices are the major determinants of demand according to the traditional theory of demand;

Second, individual demand is different from market demand, and for market demand to hold, individual demand must be summed together;

Third, as income increases, proportion of income spent on inferior good tends to reduce and vice versa

Fourth, when two goods are used together, and the use of one does not reduce the use of another, those goods are tagged complimentary goods; and if otherwise, they are called substitutes.

Fifth, consumer surplus occurs whenever a consumer pays below the price, he or she is willing to pay for a good.

## 1.5 Summary

We discussed the theory of demand where determinants of demand were identified. Income, taste, price of a commodity and other prices the major determinants of demand according to the traditional theory of demand. In order to analyse market demand, individual demand was considered, and we can infer that the summation of all individual demand gives the market demand. Income consumption curve and the indifference curve were used to explain inferior and normal goods, and it was confirmed that as the household income increases, consumption of inferior goods decreases, thus leading to an increase in the consumption of a normal good. Concept of substitution and complementary goods were discussed. It was concluded that if an increase in the price of one good leads to a fall in the demand of another good, the goods concerned are complementary goods, otherwise substitute. Finally, consumer surplus was discussed, and the demand curve diagram was used to identify consumer surplus position. We can infer that the consumer surplus position lies above the supply curve and equilibrium, but below the demand curve. This simply means the consumer is paying less than what he was willing to pay *ab initio*.

## 1.6 Tutor-Marked Assignment

Submit an essay of two pages (A4, 1.5 spacing, 12pts, Times New Roman Font) on the theory of demand, analysis of consumer surplus using demand curve diagram, inferior and normal good using Income- Consumption curve and the Indifference Curve.

## 1.7 References/Further Reading

Peter, E. & Peter, C. (1990). *Principles of Microeconomics.* Unwin Hyman.

Robert, S. P. & Daniel, L. R. (2009). *Microeconomics.* (7th ed.). New Jersey: Pearson Education International.

Stephen, H. (1999). *Applied Microeconomics.* Wiley.

Sexton, R. L. (2013). *Exploring Economics*. (6th ed.). Mason, OH: South-Western Cengage Learning

### 1.8 Answers to Self-Assessment Exercises

**Question**

Describe the inferior and normal goods using indifference curve and income consumption curve

**Model Answer**

Indifference Curve Approach to Normal and Inferior Goods

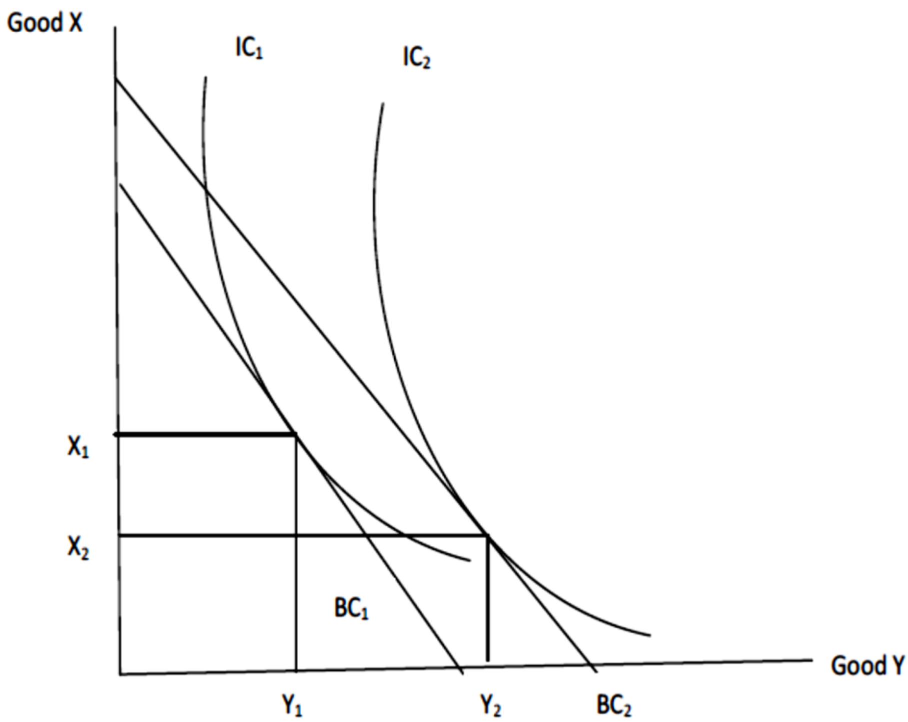



**Fig. 2.4: Indifference Curve** (IC = Indifference Curve BC = Budget Constraint)

Depending on indifference curve, the amount of a good bought can either increase, decrease, or stay the same when income increases. In Figure 2.4, good Y is a normal good since the amount purchased increases from $Y_1$ to $Y_2$ as the budget constraint shifts from $BC_1$ to the higher income $BC_2$. Good X is an inferior good since the amount bought decreases from $X_1$ to $X_2$ as income increases.

**Question**

Explain the concept of consumer surplus with graphical illustration.

Consumer surplus is the monetary gain obtained by consumers because they are able to purchase a product for a price that is less than the highest price that they would be willing to pay. The concept of consumer surplus was introduced to deal with novel complexity raised by the dominance of monopoly capital. On a supply and demand diagram, consumer surplus is the area above the equilibrium price of the good and below the demand curve. This reflects the fact that consumers would have been willing to buy a single unit of the good at a price higher than the equilibrium price, a second unit at a price below that but still above equilibrium, etc., yet they in fact pay just the equilibrium price for each unit they buy.

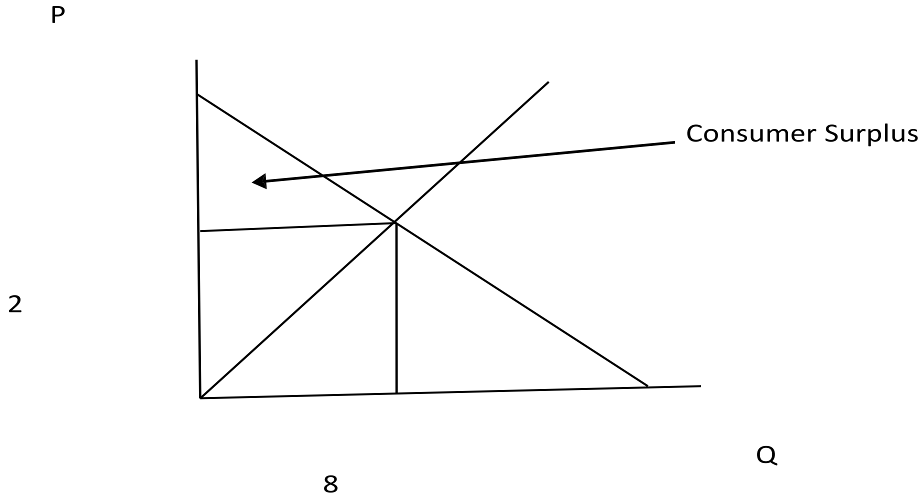



**Figure 2.5: Consumer Surplus**

The maximum amount a consumer would be willing to pay for a given quantity of a good is the sum of the maximum price he would be willing to pay for the first unit, the (lower) maximum price he would be willing to pay for the second unit, etc.

For a given price the consumer buys the amount for which the consumer surplus is highest

$$x2 = \frac{m}{p2} - \frac{p1}{p2}x1 \qquad (4)$$

where consumer surplus is the sum, over all units, of the excess of the maximum willingness to pay over the equilibrium (market) price. The consumer's surplus is highest at the largest number of units for which, even for the last unit, the maximum willingness to pay is not below the market price. Figure 2.5 depicts the consumer surplus.

# Unit 2 Elasticities of Demand

**Contents**



## 2.1 Introduction

Elasticity of demand is an important variation of the concept of demand. Demand can be classified as elastic or inelastic. It is the degree to which the demand for a good or service varies with its price. Normally, sales increase with drop in prices and decrease with rise in prices. As a general rule, appliances, cars, confectionary and other non-essentials show elasticity of demand whereas most necessities (food, medicine, basic clothing) show inelasticity of demand (do not sell significantly more or less with changes in price).

An elastic demand is one in which the change in quantity demanded due to a change in price is large. An inelastic demand is one in which the change in quantity demanded due to a change in price is small. The formula for computing elasticity of demand is:

$$\frac{(Q1 - Q2)}{(Q1 + Q2)} \div \frac{(P1 - P2)}{(P1 + P2)}$$

If the formula creates a number greater than 1, the demand is elastic. In other words, quantity changes faster than price. If the number is less than 1, demand is inelastic. In other words, quantity changes slower than price. If the number is equal to 1, elasticity of demand is unitary. In other words, quantity changes at the same rate as price.

## 2.2 Objectives

At the end of this unit, you should be able to:

- define and calculate elasticity of demand
- define and calculate the price elasticity of demand and its effect on normal and inferior goods
- have an in-depth knowledge of the term arc elasticity of demand and cross elasticity of demand.

## 2.3 Main Content

### 2.3.1 Elastic Demand

Elasticity of demand is illustrated in Figure 2.6. Note that a change in price results in a large change in quantity demanded. An example of products with an elastic demand is consumer durable which is depicted in Figure 2.6. These are items that are purchased infrequently, like a washing machine or an automobile, and can be postponed if price rises. For example, automobile rebates have been very successful in increasing automobile sales by reducing price. Close substitutes for a product affect the elasticity of demand. If another product can easily be substituted for your product, consumers will quickly switch to the other product if the price of your product rises or the price of the other product declines. For example, beef, pork and poultry are all meat products. The declining price of poultry in recent years has caused the consumption of poultry to increase, at the expense of beef and pork. So, products with close substitutes tend to have elastic demand.

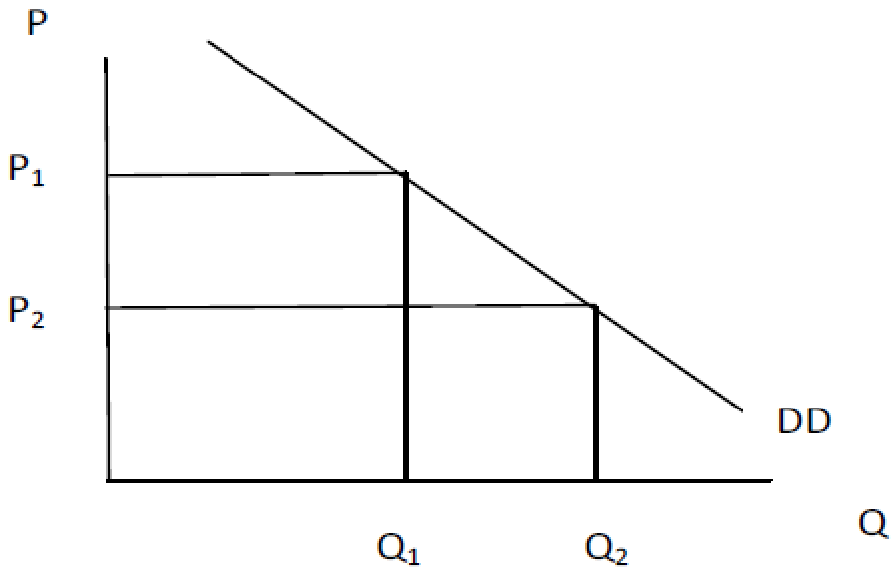



**Fig. 2.6: Elastic Demand**

For example, if $P_1$ = ₦10

$P_2$ = ₦8

$Q_1 = 30$

$Q_2 = 50$

Elasticity using the above formula = $\frac{(50-30)}{(50+30)} \div \frac{(10-8)}{(10+8)}$ = 20 ÷ 80/ 2 ÷ 18

Elasticity = 9/4 = 2.25.

**Self-Assessment Exercise**

How can you determine if demand is elastic?

### 2.3.2 Inelastic Demand

Inelastic demand is shown in Figure 2.7. Note that a change in price results in only a small change in quantity demanded. In other words, the quantity demanded is not very responsive to changes in price. Examples of this are necessities like food and fuel; consumers will not reduce their food purchases if food prices rise, although there may be shifts in the types of food they purchase. Also, consumers will not greatly change their driving behaviour if gasoline prices rise.

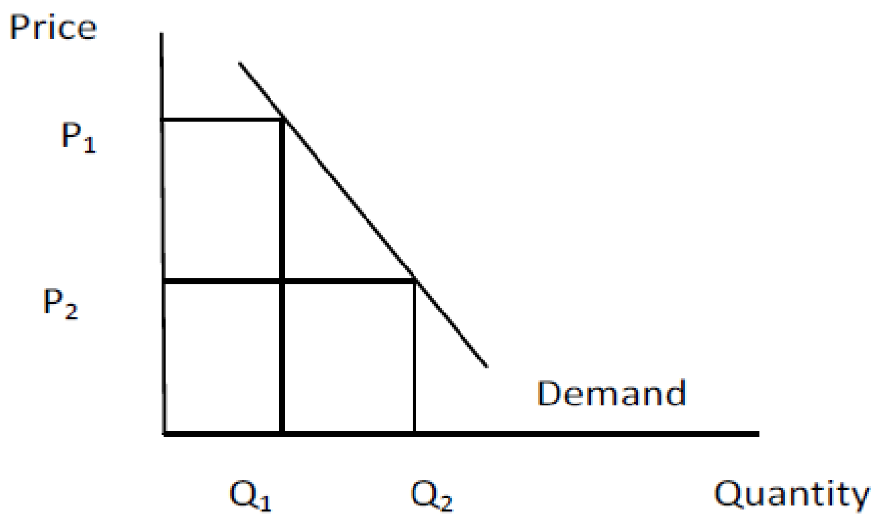



**Fig. 2.7: Inelastic Demand**

For example, assuming:

P1 = ₦12

P2 = ₦6

Q1 = 40

Q2 = 50

$$\frac{(50-40)}{(50+40)} \div \frac{(12-6)}{(12+6)} = 10 \div 90 \,/\, 6 \div 18$$

This does not mean that the demand for an individual producer is inelastic. For example, a rise in the price of petroleum at all stations may not reduce petroleum sales significantly. However, a rise of an individual station's price will significantly affect that station's sales.

**Self-Assessment Exercise**

How do we determine inelastic demand?

## 2.3.3 Price or Point Elasticity of Demand

This is a measure of responsiveness of the quantity of a raw good or service demanded to changes in its price. The formula for the price elasticity of demand is:

PED = (% change in Quantity demanded) / (% change in Price).

For example, if the price of a mobile phone changes from ₦9 to ₦10, and quantity bought changes from 150 to 110, calculate the price elasticity of demand.

$P_0$ = ₦9

$P_1$ = ₦10

$Q_0 = 150$

$Q_1 = 110$

Where $P_0$ and $Q_0$ represent the old price and quantity and $P_1$ and $Q_1$ represent the new price and quantity. To calculate the price elasticity, we need to know what the percentage change in quantity demand is and what the percentage change in price is. It's best to calculate this one at a time.

$Q_1 – Q_0/Q_0$ = Percentage change in quantity. (110-150) / 150 = -0.2667 (1)

The above figure is the percentage change in quantity demanded. We need to also calculate the percentage change in price.

$P_1 – P_0/P_0$ = Percentage change in price

(10 – 9) / 9 = 0.1111 (2)

Applying the (1) and (2) to the PED formula gives the PED figure. PED = (-0.2667) / (0.1111) = -2.4005.

In analysing price elasticities, we are concerned with their absolute value, so we ignore the negative sign. Based on the final outcome of the PED, conclusion can be drawn that the price elasticity of demand when the price increases from ₦9 to ₦10 is 2.4005.

**Interpretation of Price Elasticity of Demand**

If PED > 1, then demand is price elastic (demand is sensitive to price changes)

If PED = 1, then demand is unit elastic

If PED < 1, then demand is price inelastic (demand is not sensitive to price changes)

**Self-Assessment Exercise**

What are the decision rules for price elasticity of demand?

### 2.3.4 Arc Elasticity of Demand

The arc elasticity is the elasticity of one variable with respect to another between two given points. It is the ratio of the percentage change of one of the variables between the two points to the percentage change of the other variable. It contrasts with the point or price elasticity, which is the limit of the arc elasticity, as the distance between the two points approaches zero and which hence, is defined at a single point rather for a pair of points.

This approach is used when there is no general function to define the relationship of the two variables. Arc elasticity of demand is the percentage change in quantity demanded for a unit change in price. Arc elasticity computes the percentage change between two points in relation to the average of the two prices and the average of the two quantities, rather than the change from one point to the next.

The **y** arc elasticity of **x** is defined as:

$$E_{x,y} = \frac{\%\ change\ in\ x}{\%\ change\ in\ y}$$

Where the percentage change in going from point 1 to point 2 is usually calculated relative to the midpoint:

$$\%\text{ change in x} = \frac{(x2 - x1)}{(x2 + x1)/2}$$

$$\%\text{ change in y} = \frac{(y2 - y1)}{(y2 + y1)/2}$$

The arc elasticity of quantity demanded Q with respect to price P is broadly calculated as:

(% change in Q) / (% change in P).

For example, suppose that the two points on a demand curve, ($Q_1$, $P_1$) and ($Q_2$, $P_2$) are known, the arc elasticity is obtained using the formula:

$$E_p = \frac{\frac{Q2 - Q1}{(Q1+Q2)/2}}{\frac{P2 - P1}{(P1 + P2)/2}}$$

Suppose the quantity of meat-pie demanded at halftime of football games is measured at two different games: at one measurement it is 80 units, and at the other measurement it is 120 units. The percent change, measured against the average, would be (120-80)/((120+80)/2))=40%.

In contrast, if the percentage change were measured against the initial value, the result would be (120-80)/80= 50%. The percent change for the opposite trend, 120 units to 80 units, would be (80-120)/120 = -33.3%. The midpoint formula has the benefit that a movement from A to B is measured as the same as a movement from B to A in absolute value. (In this case, the movement from 120 to 80 would be measured as -40
%.)

Suppose that the change in the price of meat-pie, which led to this change in quantity demanded from 80 to 120, was from ₦3 to ₦1. The percent change in price measured against the midpoint would be (1-3)/2 = -100%, so the price elasticity of demand is 40%/ (-100%) or -40%. It is common to refer to the absolute value of the price elasticity as simply price elasticity, since for a normal (decreasing) demand curve the elasticity is always negative and so the “minus” part can be made implicit. Thus, the demand of the football fans for meat-pie has 40% elasticity.

**Self-Assessment Exercise**

Why do we ignore the negative sign of arc elasticity of demand?

## 2.3.5 Income Elasticity of Demand

This is a measure of the relationship between a change in the quantity demanded for a particular good and a change in real income. Put differently, income elasticity of demand measures the responsiveness of the demand for a good to a change in the income of the people demanding the good, ceteris paribus. It is calculated as the ratio of the percentage change in demand to the percentage change in income, i.e., if in response to a 10% increase in income, the demand for a good increased by 20%, the income elasticity of demand would be 20%/10% = 2.

Income elasticity of demand is calculated basically using this formula: IED = (% change in quantity demanded) / (% change in income).

Consider another example where the quantity demanded of a good increase by 15% due to a 10% increase in income, the income elasticity of demand would be 15% / 10% = 1.5. The degree to which the quantity demanded for good changes in response to a change in income depends on whether the good is a necessity or a luxury.

Normal good have positive income elasticity of demand. As income rises, more goods are demanded at each price level. The quantity demanded for normal necessities will increase with income, but at a slower rate than luxury goods. This is because consumers, rather than buying more of the necessities, will likely use their increased income to purchase more luxury goods and services. During a period of increasing incomes, the quantity demanded for luxury products tends to increase at a higher rate than the quantity demanded for necessities. The quantity demanded for luxury goods is very sensitive to changes in income. On the other hand, inferior goods have a negative income elasticity of demand - the quantity demanded for inferior goods falls as income rise. For example, the quantity demanded for generic food items tends to decrease during periods of increased incomes.

**Possible outcome of income elasticity of demand**

(i) High income elasticity of demand: In this case, increase in income is accompanied by relatively larger increase in quantity demanded. Here the value of coefficient Ey is greater than unity (Ey>1). E.g.: 20% increase in quantity demanded due to 10% increase in income.

(ii) Unitary income elasticity of demand: In this case increase in income is accompanied by equal proportionate increase in quantity demanded. Here the value of coefficient Ey is equal to unity (Ey=1). E.g.: 10% increase in quantity demanded due to 10% increase in income.

(iii) Low income elasticity of demand: In this case proportionate increase in income is accompanied by relatively lower increase in quantity demanded. Here the value of coefficient Ey is less than unity (Ey<1). E.g.: 5% increase in quantity demanded due to 10% increase in income.

(iv) Zero income elasticity of demand: This shows that quantity bought is constant regardless of changes in income. Here the value of coefficient Ey is equal to zero (Ey=0). E.g.: No change in quantity demanded even when there is 10% increase in income.

(v) Negative income elasticity of demand: In this case increase in income is accompanied by decrease in quantity demanded. Here the value of coefficient Ey is less than zero/negative (Ey<0). E.g.: 5% decrease in quantity demanded due to 10% increase in income.

**Self-Assessment Exercise**

What is the formula for calculating income elasticity of demand?

### 2.3.6 Cross Elasticity of Demand

Cross elasticity of demand measures the responsiveness of the demand for a good to a change in the price of another good. It is measured as the percentage change in demand for the first good that occurs in response to a percentage change in price of the second good. For example, if in response to a 10% increase in the price of fuel, the demand for new cars that are fuel sufficient decreased by 20%, the cross elasticity of demand would be: -20% / 10% = -2.

A negative cross elasticity denotes two products that are complements, while a positive cross elasticity denotes two substitute products. Assuming products A and B are complements, meaning that an increase in the demand for A is caused by an increase in the quantity demanded for B. Therefore, if the price of product B decreases, then the demand for product A shifts to the right, increasing A's demand, resulting in a negative value for the cross elasticity of demand. The exact opposite reasoning holds for substitutes.

$$\frac{\%\ change\ in\ quantity\ demanded\ of\ product\ A}{\%\ change\ in\ price\ of\ product\ B}$$

Or

$$Ec = \frac{P1A + P2A}{Q1B + Q2B} \times \frac{\Delta QB}{\Delta PA}$$

Where:

P1A = The price of good A at time period 1

P2A = The price of good A at time period 2

Q1B = The quantity demanded of good B at the time period 1

Q2B = The quantity demanded of good B at the time period 2

ΔQB = The change in the quantity demanded of good B ΔPA = The change in price of good A

The cross elasticity of demand for substitute goods will always be positive, because the demand for one good will increase if the price for the other good increases. For example, if the price of coffee increases (but everything else stays the same), the quantity demanded for tea (a substitute beverage) will increase as consumers switch to an alternative. On the other hand, the coefficient for complements will be negative. For example, if the price of coffee increases (but everything else stays the same), the quantity demanded for coffee stir sticks will drop as consumers will purchase fewer sticks. If the coefficient is 0, then the two goods are not related.

**Self-Assessment Exercise**

Define the term cross elasticity of demand.

## 2.4 Conclusion

From our discussion so far on the elasticity of demand, we can infer the following:

First, when an increase in price leads to a decrease in quantity demanded, such demand can be referred to as inelastic otherwise elastic.

Second, percentage change in quantity demanded of product A divided by the percentage change in price of product B gives the cross elasticity of demand

Third, income elasticity of demand is calculated as the percentage change in quantity demanded divided by percentage change in income.

## 2.5 Summary

Elasticity of demand is classified into elastic and inelastic. When the change in price results in a large change in quantity demand; such demand is elastic otherwise inelastic. Price elasticity of demand measures the responsiveness of the quantity demanded of a good to changes in its price. Arc elasticity of demand measures the elasticity of one variable with respect to another between two given points. Income elasticity of demand. This measures the relationship between a change in the quantity demanded for a particular good and a change in real income. Finally, the cross elasticity of demand measures the responsiveness of the demand for a good to a change in the price of another good.

## 2.6 Tutor-Marked Assignment

1. Assuming an increase in price from ₦10 to ₦15 led to a decrease in the quantity demanded from 20 to 12.
   (a) Calculate the elasticity of demand and interpret the result. (b) Calculate the price elasticity of demand
2. $P_1$ = ₦5, $P_2$ = ₦7

   $Q_1 = 4$, $Q_2 = 2$

Using the above information, calculate:

(a) Arc elasticity of demand

(b) Cross elasticity of demand.

## 2.7 Reference/Further Reading

Koutsoyiannis, A. (1987). *Modern Microeconomics*. (2nd ed.). International.

McEachern, W. A. (2014). *Economics: A Contemporary Introduction*. (10th ed.). Mason, OH: South-Western Cengage Learning.

Miller, R. L. (2014). *Economics Today*. (17th ed.). Upper Saddle River, NJ: Pearson.

O'Sullivan, et *al*. (2014). Economics*: Principles, Applications, and Tools*. (8th ed.). Upper Saddle River, N. J.: Pearson.

## 2.8 Answers to Self-Assessment Exercises

Why do we ignore the negative sign of arc elasticity of demand?

**Model Answer**

It is common to refer to the absolute value of the price elasticity as simply price elasticity. This is because for a normal (decreasing) demand curve the elasticity is always negative. Consequently, the "minus" part is made implicit. Implicit in the sense that the elasticity value is considered in the absolute sense, internalizing the decreasing nature of a normal demand curve.

## Unit 3 Budget Constraint

**CONTENTS**



### 3.1 Introduction

A budget is a quantitative expression of a plan for a defined period of time. It may include planned sales volumes and revenues, resource quantities, costs and expenses, assets, liabilities and cash flows. It expresses strategic plans of business units, organisations, activities or events in measurable terms. Budget constraint represents all the combinations of goods and services that a consumer may purchase given current prices within his or her given income.

### 3.2 Objectives

At the end of this unit, the student should be able to:

- define and analyse the term budget constraint
- explain the effect of price and income changes on the budget line
- analyse how the budget line changes
- describe the combined effects of price and subsidy on consumer income.

## 3.3 Main Content

### 3.3.1 The Concept of Budget

Budget is an estimation of the revenue and expenses over a specified future period of time. A budget can be made for a person, family, group of people, business, government, country, multinational organisation or just about anything else that makes and spends money. A budget in microeconomic is a concept that shows the tradeoff made when one good is exchanged for another. In other words, a budget is an important concept in microeconomics, which uses a budget line to illustrate the trade-offs between two or more goods.

Budget aids the planning of actual operations by forcing producers to consider how the conditions might change and what steps should be taken now and by encouraging managers to consider problems before they arise. It also helps co-ordinate the activities of the organisation by compelling managers to examine relationships between their own operations and those of other departments.

**Self-Assessment Exercise**

Define the term budget

### 3.3.2 The Concept of Budget Constraint

A budget constraint is an accounting identity that describes the consumption options available to a consumer with a limited income (or wealth) to allocate among various goods. It is important to understand that the budget constraint is an accounting identity, not a behavioural relationship. A consumer may well determine his behaviour by considering his budget constraint, but his budget constraint is a given element of the problem he faces. For instance, suppose that there is some set of goods from which the consumer can choose. In real life there are many goods to consume, but for our purposes it is convenient to consider only the case of two goods. We will indicate the consumer's consumption bundle by $(x_1, x_2)$. This is simply a list of two numbers that tells us how much the consumer is choosing to consume of good 1, $x_1$, and how much the consumer is choosing to consume of good 2, $x_2$.

We suppose that we can observe the prices of the two goods, ($p_l$ and $p_2$), and the amount of money the consumer has to spend, m. Then the budget constraint of the consumer can be written as:

$$p1x1 + p2x2 \le m \qquad (1)$$

Here, p1x1 is the amount of money the consumer is spending on Good 1, and $p_2x_2$ is the amount of money the consumer is spending on Good 2.

The budget constraint of the consumer requires that the amount of money spent on the two goods be no more than the total amount the consumer has to spend. The consumer's affordable consumption bundles are those that don't cost any more than m. We call this set of affordable consumption bundles at prices ($p_l$, $p_2$) and income m the budget set of the consumer.

The two-good assumption is more general, since we can often interpret one of the goods as representing everything else the consumer might want to consume. For example, if we are interested in studying a consumer's demand for milk, we might let $x_l$ measure his or her consumption of milk in tins per month. We can then let $x_2$ stand for everything else the consumer might want to consume.

When we adopt this interpretation, it is convenient to think of good 2 as being the naira that the consumer can use to spend on other goods. Under this interpretation the price of good 2 will automatically be 1, since the price of one naira is one naira. Thus, the budget constraint will take the form:

$$p1x1 + x2 \le m \qquad (2)$$

This expression simply says that the amount of money spent on Good 1, $p_1x_1$, plus the amount of money spent on all other Goods, $x_2$, must be no more than the total amount of money the consumer has to spend, m.

We say that Good 2 represents a composite good that stands for everything else that the consumer might want to consume other than Good 1. Such a composite good is invariably measured in naira to be spent on goods other than Good 1. As far as the algebraic form of the budget constraint is concerned, equation (2) is just a special case of the formula given in equation (1), with $p_2 = 1$, so everything that we have to say about the budget constraint in general will hold under the composite good interpretation.

**Self-Assessment Exercise**

i. What is budget constraint?

ii. Express budget constraint algebraically.

### 3.3.3 Properties of the Budget Set

Budget set consist of all bundles that are affordable at the given prices and income. The budget line is the set of bundles that cost exactly m:

$$p1x1 + p2x2 = m \tag{3}$$

These are the bundles of goods that just exhaust the consumer's income. The budget set is depicted in diagram below. The downward sloping line is the budget line which also represents the bundles that cost exactly m, and the bundles below this line are those that cost strictly less than m.

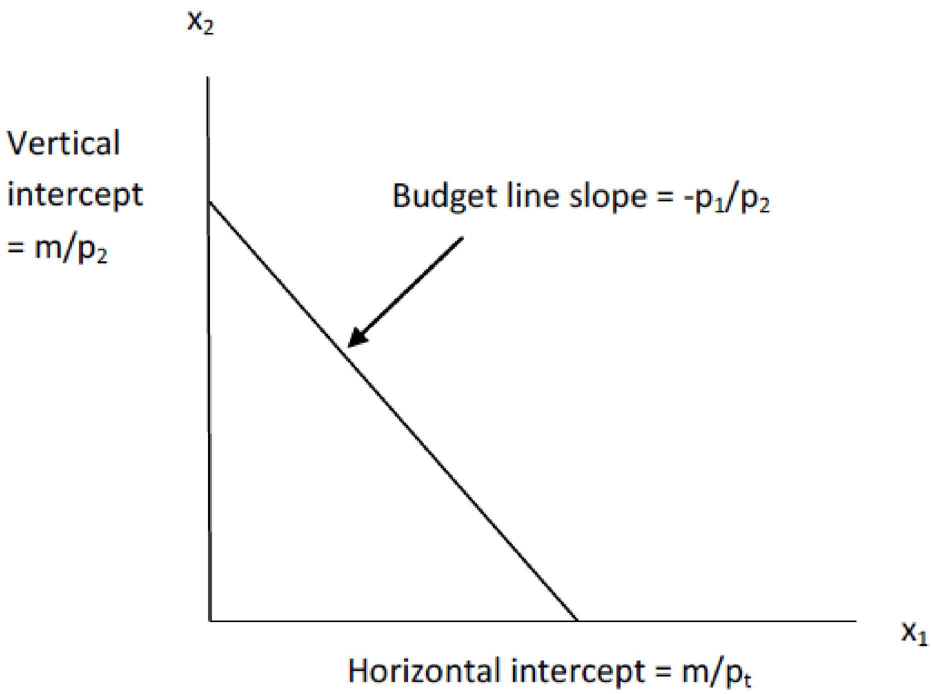



**Fig. 2.7a: Budget Line Analysis**

Rearranging the budget line in equation (3) gives us the formula:

$$x2 = \frac{m}{p2} - \frac{p1}{p2}x1 \tag{4}$$

The above formula tells us how many units of Good 2 the consumer needs to consume in order to just satisfy the budget constraint if she is consuming xl units of Good 1.

Here is an easy way to draw a budget line given prices ($p_l$ , $p_2$) and income m. The question of how much of Good 2 the consumer could buy if she spent all of her money on Good 2 is then asked. The answer is, of course, $m/p_2$. Then ask how much of Good 1 the consumer could buy if she spent all of her money on Good 1. The answer is $m/p_l$. Thus, the horizontal and vertical intercepts measure how much the consumer could get if she spent all of her money on Goods 1 and 2, respectively.

The slope of the budget line has economic interpretation. It measures the rate at which the market is willing to "substitute" Good l for Good 2. Suppose for example that the consumer is going to increase her consumption of Good 1 by $\Delta x_l$.' How much will her consumption of Good 2 have to change in order to satisfy her budget constraint? Let us use $\Delta x_2$ to indicate her change in the consumption of Good 2. Now note that if she satisfies her budget constraint before and after making the change she must satisfy:

$$p1x1 + p2x2 = m$$

And

$$p1(x1 + \Delta x1) + p2(x2 + \Delta x2) = m$$

Subtracting the first equation from the second gives:

$$p1\,\Delta x1 + p2\Delta x2 = 0$$

This says that the total value of the change in her consumption must be zero. Solving for $\Delta x_2/\Delta x_1$, the rate at which Good 2 can be substituted for Good 1 while still satisfying the budget constraint, gives:

$$\frac{\Delta x2}{\Delta x1} = -\frac{p1}{p2}$$

This is just the slope of the budget line. The negative sign on the price ratio is there since $\Delta x_1$ and $\Delta x_2$ must always have opposite signs. The negative sign on the ratio of the prices of good x1 and Good x2 implies that if you consume more of Good 1, you have to consume less of Good 2 and vice versa for you to continue to satisfy the budget constraint.

**Self-Assessment Exercise**

Identify the properties of a budget line

### 3.3.4 How the Budget Changes

When prices and incomes change, the set of goods that a consumer can afford changes as well. To determine how these changes, affect the budget set, we will have to consider changes in income. It is easy to see from equation (4) that an increase in income will increase the vertical intercept and not affect the slope of the line. Thus, an increase in income will result in a parallel shift outward of the budget line as in Figure 2.7b. Similarly, a decrease in income will cause a parallel shift inward.

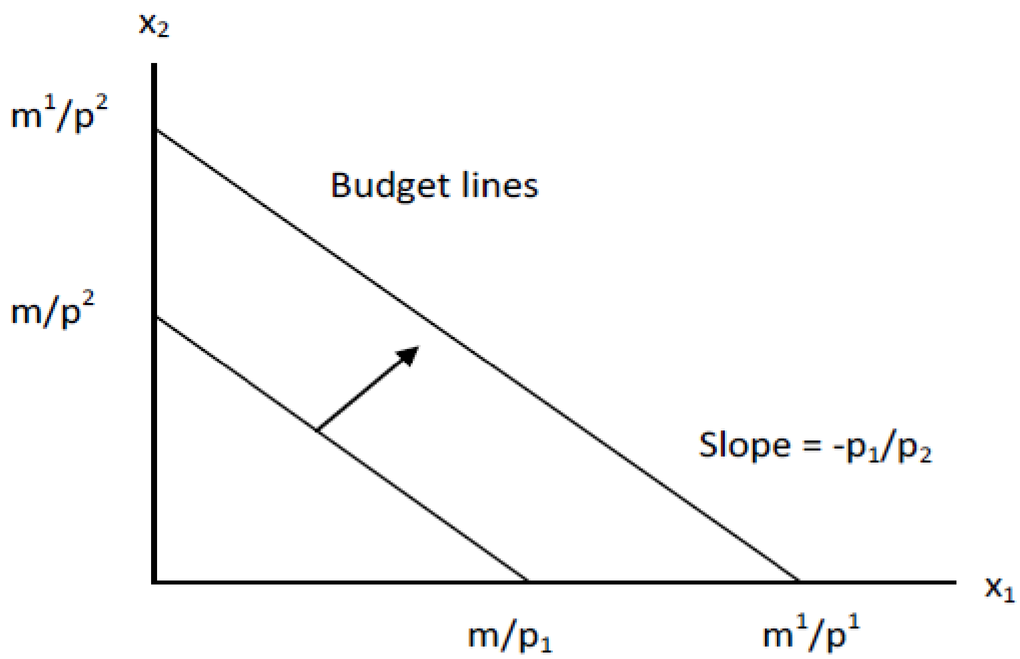



**Fig. 2.7b: Budget Line Analysis**

We need to consider changes in price as well, in order to do this, let us first consider increasing price 1 while holding price 2 and income fixed. According to equation (4), increasing $p_l$ will not change the vertical intercept, but it will make the budget line steeper since $p_l/p_2$ will become larger.

Another way to see how the budget line changes is to use the trick described earlier for drawing the budget line. If you are spending all of your money on Good 2, then increasing the price of Good 1 does not change the maximum amount of Good 2 you could buy, thus the vertical intercept of the budget line does not change. But if you are spending all of your money on Good 1, and Good 1 becomes more expensive, then your consumption of Good 1 must decrease. Thus, the horizontal intercept of the budget line must shift inward, resulting in the Figure 2.7c.

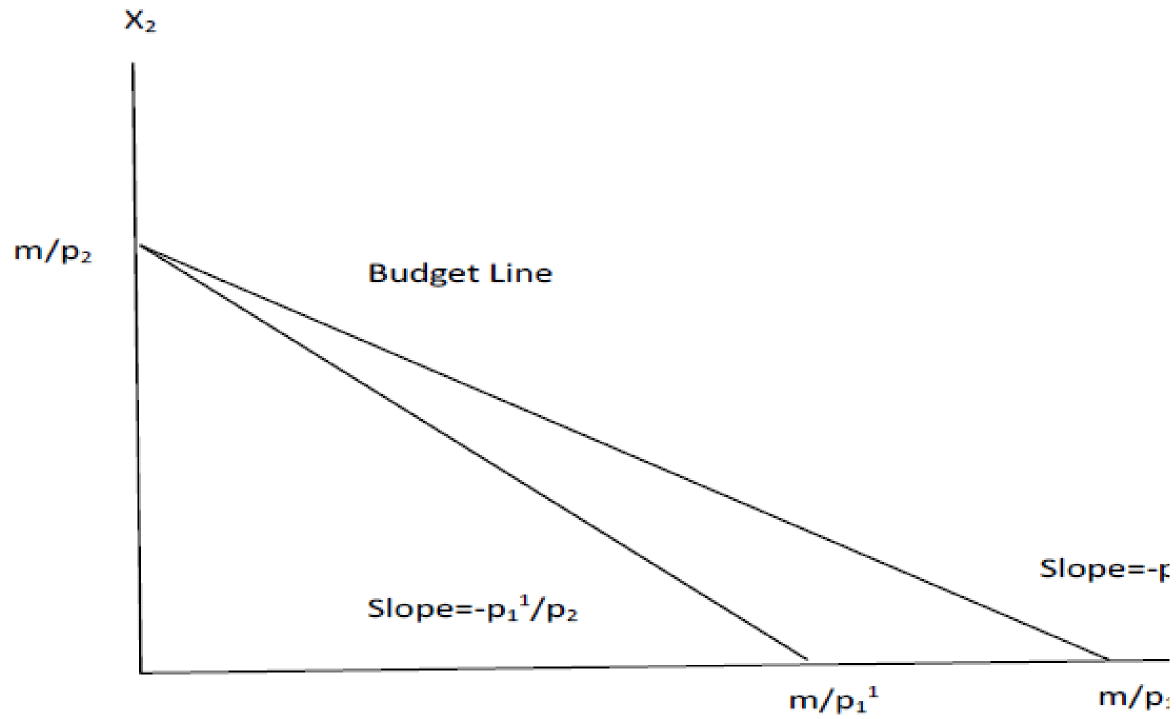



**Fig. 2.7c: Budget Line Analysis**

What happens to the budget line when the prices of Good 1 and Good 2 change at the same time? Suppose for example that we double the prices of both Goods 1 and 2. In this case, both the horizontal and vertical intercepts shift inward by a factor of one-half, and therefore the budget line shifts inward by one-half as well. Multiplying both prices by two is just like dividing income by 2. We can also see this algebraically. Suppose our original budget line is:

$$p1x1 + p2x2 = m$$

Now suppose that both prices become *t* times as large. Multiplying both prices by *t* yields:

$$tp1x1 + tp2x2 = m$$

But this equation is the same as:

$$p1x1 + tp2x2 = \frac{m}{t}$$

Thus, multiplying both prices by a constant amount t is just like dividing income by the same constant *t*. It follows that if we multiply both prices by *t* and we multiply income by *t*, then the budget line will not change at all. We can also consider price and income changes together. What happens if both prices go up and income goes down? Think about what happens to the horizontal and vertical intercepts. If m decreases and pl and p2 both increases, then the intercepts m/pl and m/p2 must both decrease. This means that the budget line will shift inward. What about the slope of the budget line? If Price 2 increases more than Price 1, so that -pl/p2 decreases (in absolute value), then the budget line will be flatter; if Price 2 increases less than Price 1, the budget line will be steeper.

**Self-Assessment Exercise**

Highlight and explain the factors that cause changes in budget line.

### 3.3.5 Implication of Taxes and Subsidies on Budget

Economic policy often uses tools that affect a consumer's budget constraint, such as taxes. For example, if the government imposes a quantity tax, this means that the consumer has to pay a certain amount to the government for each unit of the good he purchases. Consider for example, a ₦15 tax on a liter of petroleum. How does this quantity tax affect the budget line of a consumer? From the viewpoint of the consumer, the tax is just like a higher price. Thus, a quantity tax of t naira per unit of Good 1 simply changes the price of Good 1 from $p_l$ to $p_l + _t$. This implies that the budget line must get steeper.

Another kind of tax is a value tax. As the name implies, this is a tax on the value-the price-of a good, rather than the quantity purchased of a good. A value tax is usually expressed in percentage terms. Consider for example if the sales tax is 6 per cent, then a good that is priced at ₦1 naira will actually sell for ₦1.06. (Value taxes are also known as ad- valorem taxes). If Good 1 has a price of pl but is subject to a sales tax at rate 7, then the actual price facing the consumer is (1 + t) pl. The consumer has to pay pl to the supplier and $tp_l$ to the government for each unit of the good. Therefore, the total cost of the good to the consumer is $(1 + t)$ pl.

**A subsidy** is the opposite of a tax. In the case of a quantity subsidy, the government gives an amount to the consumer that depends on the amount of the good purchased. If, for example, the consumption of milk was subsidised, the government would pay some amount of money to each consumer of milk depending on the amount that consumer purchased. If the subsidy is *s* naira per unit of consumption of Good 1, then from the viewpoint of the consumer, the price of Good 1 be $p_l - s$. This would therefore make the budget line flatter.

Similarly, an ad-valorem subsidy is a subsidy based on the price of the good being subsidised. If the government gives you back ₦1 for every ₦2 you donate to charity, then your donations to charity are being subsidised at a rate of 50 per cent. In general, if the price of Good 1 is $p_l$ and Good 1 is subject to an ad-valorem subsidy at rate *a*, then the actual price of Good 1 facing the consumer is $(1 - a)$ pl.

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

Governments also sometimes impose rationing constraints. This means that the level of consumption of some good is fixed to be no larger than some amount. Suppose, for example, that Good 1 were rationed so that no more than $x_1^*$ could be consumed by a given consumer. Then the budget set of the consumer would look like that depicted in the diagram below: it would be the old budget set with a piece lopped off. The lopped-off piece consists of all the consumption bundles that are affordable but have $x_l > x_1*$.

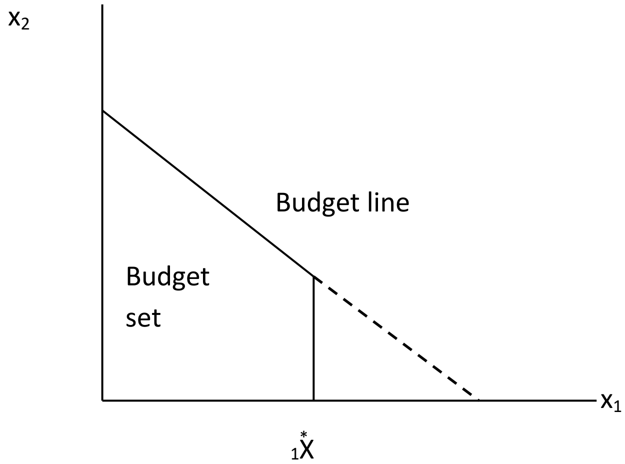



**Fig. 2.7d: Budget Line Analysis**

Sometimes taxes, subsidies, and rationing are combined. For example, we could consider a situation where a consumer could consume Good 1 at a price of p1 up to some level $x_1$, and then had to pay a tax t on all consumption in excess of $x_I$. The budget set for this consumer is depicted in Figure 2.7e. Here the budget line has a slope of $-p_l/p_2$ to the left of $x_1$, and a slope of $-(p_l + t)/p_2$ to the right of $x_1$.

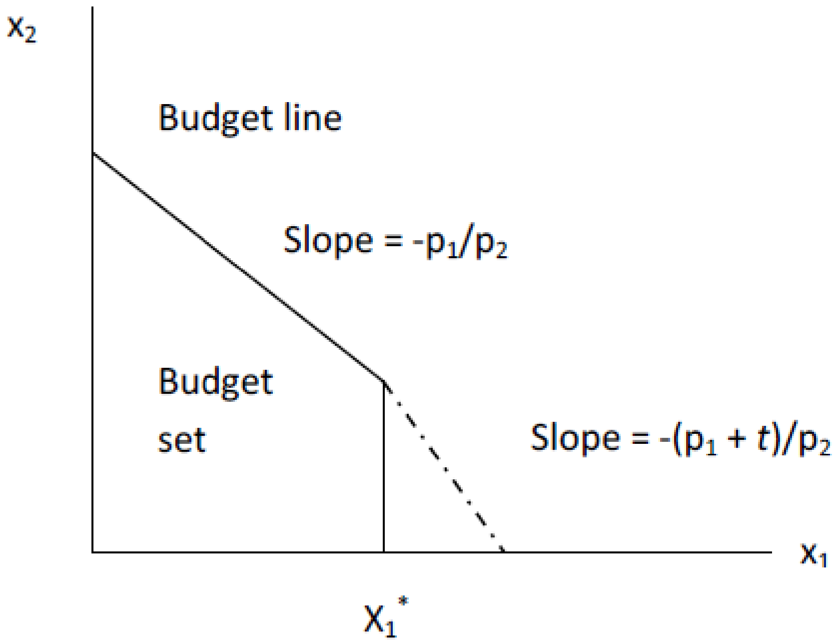



**Fig. 2.7e: Budget Line Analysis**

**Self-Assessment Exercise**

What are the effects of subsidy and taxes on the budget line?

## 3.4 Conclusion

From our discussion so far on budget constraint, we can infer the following:

- First, that the amount of money consumers can spend on goods cannot be more than his income;
- Second, $p_1x_1+p_2x_2=m$ represents the budget constraint, that is, the prices and goods that money income (m) can buy at a point in time
- Third, when prices and incomes change, the set of goods that a consumer can afford changes as well
- Fourth, when the price of Good A increases while price of Good B and income remain constant, the budget line becomes steeper and vice versa
- Fifth, when both prices of Good A and B go up while income reduces, the budget line shifts inward.

## 3.5 Summary

The budget set consists of all bundles of goods that the consumer can afford at given prices and income. We assumed that there are only two goods, but this assumption is more general than it seems. The budget line is written as $p_lx_l + p_2x_2 = m$. It has a slope of $-p_l/p_2$, a vertical intercept of $m/p_2$, and a horizontal intercept of m/pl. Increasing income shifts the budget line outward. Increasing the price of good 1 makes the budget line steeper. Increasing the price of good 2 makes the budget line flatter. Taxes, subsidies, and rationing change the slope and position of the budget line by changing the prices paid by the consumer.

## 3.6 Tutor-Marked Assignment

1. If the income of the consumer increases and one of the prices decreases at the same time, will the consumer necessarily be at least well-off?
2. If the price of Good 1 doubles and the price of Good 2 triples do the budget line become flatter or steeper?
3. Suppose that the government puts a tax of ₦20 naira on a liter of petroleum and then later decides to put a subsidy on petroleum at a rate of ₦12 naira per liter. What net tax is this combination equivalent to?
4. What happens to the budget line if the price of Good 2 increases, but the price of Good 1 and income remain constant?
5. Originally the consumer faces the budget line $p_lx_l + p_2x_2 = m$. Then the price of Good 1 doubles, the price of Good 2 becomes 8 times larger, and income becomes 4 times larger. Write down an equation for the new budget line in terms of the original prices and income.

## 3.7 References/Further Reading

Cowen, T. & Alex, T. (2012). *Modern Principles: Economics*. (2nd ed.). New York: Worth Publishers.

Frank, R. H. & Bernanke, B. (2013). *Principles of Economics*. (5th ed.). New York: McGraw-Hill.

Gottheil, F. M. (2014). *Principles of Economics*. (7th ed.). Mason, OH: South-Western Cengage Learning.

Koutsoyiannis, A. (1987). *Modern Microeconomics*. (International ed.).

Varian, H R. (2000). *Intermediate Microeconomics: A Modern Approach.* (7th ed.).

**3.8 Answers to Self-Assessment Exercises**

**Question**

Identify the properties of a budget line

**Model Answer**

First, the budget line is the set of bundles that cost exactly m:

$$p1x1 + p2x2 = m$$

These are the bundles of goods that just exhaust the consumer's income.

Second, rearranging the budget line equation gives us the formula:

$$x2 = \frac{m}{p2} - \frac{p1}{p2}x1 \qquad (4)$$

The above formula tells us how many units of Good 2 the consumer needs to consume in order to just satisfy the budget constraint if she is consuming xl units of Good 1.

Third, there is an easy way to draw a budget line given prices (pl , p2) and income m. if the consumer spends all of her income on Good 2, she would buy m/p2. Likewise, the consumer would buy m/pl if she spends all of her income on Good 1.

Fourth, the slope of the budget line has economic interpretation. It measures the rate at which the market is willing to "substitute" Good l for Good 2. The total value of the change in consumption must be zero.

**Question**

What are the effects of subsidy and taxes on the budget line?

**Model Answer**

From the viewpoint of the consumer, the tax is just like a higher price. Thus, a quantity tax of t naira per unit of Good 1 simply changes the price of Good 1 from $p_l$ to $p_l + _t$. This implies that the budget line must get steeper.

On the other hand, a subsidy is the opposite of a tax. In the case of a quantity subsidy, the government gives an amount to the consumer that depends on the amount of the good purchased. If, for example, the consumption of milk was subsidised, the government would pay some amount of money to each consumer of milk depending on the amount that consumer purchased. If the subsidy is *s* naira per unit of consumption of Good 1, then from the viewpoint of the consumer, the price of Good 1 be $p_l$ - *s*. This would therefore make the budget line flatter.

# Unit 4 Theory of Utility Maximization

**Contents**



## 4.1 Introduction

Utility is the ability of a good to satisfy needs or wants. Utility is also a representation of preferences over some set of goods and services. Preferences have a (continuous) utility representation so long as they are transitive, complete, and continuous. Utility is an important concept in economics and game theory, because it represents satisfaction experienced by the consumer of a good. Not coincidently, a good is something that satisfies human wants and provides utility, for example, to a consumer making a purchase. It was recognized that one cannot directly measure benefit, satisfaction or happiness from a good or service, so instead economists have devised ways of representing and measuring utility in terms of economic choices that can be counted.

Economists have attempted to perfect highly abstract methods of comparing utilities by observing and calculating economic choices. In the simplest sense, economists consider utility to be revealed in people's willingness to pay different amounts for different goods. There are two basic approaches to problem of comparison of utilities, the cardinalist approach and the ordinalist approach.

## 4.2 Objectives

At the end of this unit, you should be able to:

- differentiate between cardinal and ordinal utility theory
- explain the assumptions of Revealed Preference Hypothesis and indifference curves
- analyse the Revealed Preference Hypothesis and indifference curves
- describe the Diamond-Water paradox
- explain optimisation.

## 4.3 Main Content

### 4.3.1 Cardinal Utility Theory

Cardinal utility presupposes a measure of utility, or satisfaction derived from the consumption of goods and services that can be quantified using an absolute scale. Cardinal utility exists if the utility derived from consumption is measurable in the same way that other physical characteristics (like height and weight) are measured using a scale that is comparable between people. Put differently, cardinal utility theory argues that a consumer has the capacity to measure the level of satisfaction that she derives from consumption of a given quantity of a commodity.

Utility is a prejudiced or introspective notion associated with the internal perceptions and feelings of the customer. Therefore, it is certainly not possible for you to gauge the utility of a product to a shopper directly.

However, an indirect measure of utility is present in the price that is paid by the consumer for the specific commodity. The higher the price paid by the customer for a product, greater will be its utility. Price could, therefore, be a measure of the utility of a commodity. For example, if a shopper is willing to spend ₦2 naira for an apple and ₦1 naira for an orange, and then the utility of the apple to the buyer is twice that of the orange. To put it differently, the utilities of two distinct goods to one particular customer could be measured by the prices that he or she would like to pay for them. Money is, hence, a measuring rod, which is often used by the economists to determine utilities of goods.

**Assumptions**

1. Rationality: The consumer is rational. He aims at the maximisation of his utility subject to the constraint imposed by his income.
2. Cardinal utility: The utility of each commodity is measurable.
   Utility is a cardinal concept, and the most convenient measure is money (measured by the monetary unit the consumer is willing to pay for another unit of the commodity).
3. Constant marginal utility of money: This assumption necessary if money is used as the measure of utility. The standard of measurement must be constant. If the marginal utility of money changes as income increases (or decreases), the measuring-rod for utility becomes like an elastic ruler, thus inappropriate for measurement.
4. Diminishing marginal utility: The utility gained from consumption of successive units of a commodity diminishes. That is, the marginal utility of a commodity diminishes as the consumer acquires larger quantities of it. This is an axiom of diminishing marginal utility.
5. The total utility of a basket of goods depends on the quantity of the individual commodities. If there are n commodities in the bundle with quantities $x_1, x_2, \ldots, x_n$, the total utility is:

$$U = f(x_1, x_2, \ldots, x_n).$$

**Critique of the cardinal approach**

There are three basic weaknesses in the cardinalist approach. They are:

1. The assumptions of cardinal utility theory are unrealistic.
2. The satisfaction derived from various commodities cannot be measured objectively.
3. The assumption of constant utility of money is unrealistic. As income increases, the marginal utility of money changes.

**Self-Assessment Exercise**

List the assumptions of the cardinal utility theory.

**a. Total and Marginal Utility**

In order to analyse the above, we begin with a simple model of a single commodity x. The consumer can either buy x or retain his money income Y. under these conditions, the consumer is in equilibrium when the marginal utility of x is equated to its market price (Px).

$$MU_x = P_x$$

If the marginal utility of x is greater than its price, the consumer can increase his welfare by purchasing more of unit x. Similarly, if the marginal utility of unit x is less than its price, the consumer will increase his total satisfaction by cutting down the quantity of x and keeping more of his income unspent. Therefore, he attains the maximisation of his utility when $MU_x = P_x$. If there are more commodities, the condition for the equilibrium of the consumer is the equality of the ratios of the marginal utilities of the individual commodities to their prices:

$$\frac{MUx}{Px} = \frac{MUy}{Py} = \dots\dots.. = \frac{MUn}{Pn}$$

**Total Utility**

Total utility, as the term indicates, is the utility derived from all units of a commodity consumed. Suppose that a person consumes 10 oranges. In this case, the total utility is obtained by adding the utility derived from each unit of orange. In our example (Table 2.1), the total utility derived from the first six oranges is:

21 (21 = 6 + 5 + 4 + 3 + 2 + 1).

**Marginal Utility**

Marginal utility is the utility from the consumption of successive units of a commodity. To put it simply, marginal utility represents the utility derived from each unit of commodity under consideration.

Symbolically,

$MU = \Delta TU/\Delta C$ where,

$TU$ = total utility

$\Delta TU$ = change in total utility ($TU_n - TU_{n-1}$)

$C$ = consumption and $\Delta C = 1$ unit

In other words, marginal utility of $n^{th}$ unit of commodity A is the difference between the total utility of $n^{th}$ unit and the total utility of $(n-1)^{th}$ unit of the commodity.

Symbolically,

$MU_n = TU_n - TU_{n-1}$

Where,

$MU_n$ = Marginal utility of $n^{th}$ unit

$TU_n$ = Total utility of $n^{th}$ unit

$TU_{n-1}$ = Total utility of $(n-1)n^{th}$ unit

In our example (Table 2.1), the marginal utility of the $4^{th}$ orange is $MU_4 = TU_4 - TU_3 = 18 - 15 = 3$. Figure 2.8 details the path of total utility and marginal utility curves. The total utility curve rises initially and after certain stage, the curve starts declining. At this stage, the marginal utility curve enters into the negative zone.

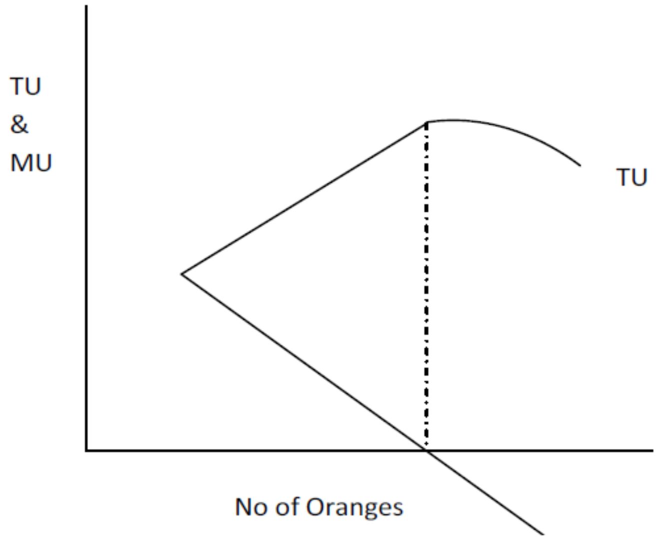



**Fig. 2.8: Total and Marginal Utility**

**Self-Assessment Exercise**

Differentiate between total and marginal utility

**b. Diminishing Marginal Utility**

The law of diminishing marginal utility states that the utility derived from each successive unit of a commodity diminishes. To put it simply, even the most beautiful place of the world or the sweetest music can make you feel bored after certain stage. The law further states that when an individual consumes more of a commodity, the total utility increases at a decreasing rate. However, after certain stage, the total utility also starts decreasing and the marginal utility becomes negative (Table 2.1). This means that the individual does not need the commodity further.

**Table 2.1: Total and Marginal Utility for Orange Consumption**

| Number of Oranges | Total Utility | Marginal Utility |
|---|---|---|
| 1 | 6 | 6 |
| 2 | 11 | 5 |
| 3 | 15 | 4 |
| 4 | 18 | 3 |
| 5 | 20 | 2 |
| 6 | 21 | 1 |
| 7 | 21 | 0 |
| 8 | 20 | -1 |

The **law of satiable** wants is also referred to as the **law of diminishing marginal utility**. The law of satiable wants postulated that an individual's want for a particular commodity gets satiated when he or she consumes more and more of it. After certain stage, the individual is not willing to consume the commodity anymore.

**Assumptions of the Law of Diminishing Marginal Utility**

**Homogeneity:** Each unit of the commodity under consideration is identical in all aspects such as quality, taste, color, size and so on.

**Reasonability:** Each unit of commodity under consideration must be same and standard. For example, 100 ml of coffee, 200 grams of apple and so on.

**Constancy:** The law of diminishing marginal utility assumes that consumer's consumption pattern, tastes, preferences, income, and price of the commodity and its substitutes are constant during the process of consumption.

**Continuity:** The law further assumes that consumption is a continuous process and there is no room for any time gap.

**Rationality:** For the law to hold well, the consumer must be a rational economic man. In addition, the law assumes that the consumer's mental condition remains normal during the process of consumption.

**Self-Assessment Exercise**
Explain briefly what you understand by the term diminishing marginal utility.

### 4.3.2 Ordinal Utility Theory

Ordinal utility theory argues that a consumer cannot measure satisfaction numerically or subjectively. Instead, she can rank the different baskets or bundles so as to choose the best basket. Put differently, the ordinalists uphold the view that amounts of utility are naturally non-measurable technically, conceptually as well as practically. They consider that the basic principles of consumer behaviour could be described without the notion of quantifiable utility. As per the idea of ordinal utility, the utilities resulting from the usage of goods can never be measured. The ordinal principle allows us to claim simply that the customer prefers an apple to an orange; however, it fails to reveal by how much. It really does not allow us to compare the value of utilities acquired from the two goods.

### a. Revealed Preference Hypothesis

Revealed preference hypothesis was introduced by Samuelson in 1938. The revealed preference hypothesis is considered a major breakthrough in the theory of demand because it has made possible the establishment of the law of demand directly (on the basis of the revealed preference axiom) without the use of indifference curves and all the restrictive assumptions.

Revealed preference hypothesis is an economic theory of consumption behaviour which asserts that the best way to measure consumer preferences is to observe their purchasing behaviour. Revealed preference theory works on the assumption that consumers have considered a set of alternatives before making a purchasing decision. Thus, given that a consumer chooses one option out of the set, this option must be the preferred option.

Revealed preference theory came about because existing theories of consumer demand were based on a diminishing marginal rate of substitution (MRS). This diminishing MRS relied on the assumption that consumers make consumption decisions to maximise their utility. While utility maximisation was not a controversial assumption, the underlying utility functions could not be measured with great certainty. Revealed Preference theory was a means to reconcile demand theory by defining utility functions and observing behaviour.

### Assumptions

**Rationality:** The consumer is assumed to behave rationally; in that he prefers bundles of goods that include more quantities of the commodities.

**Consistency:** The consumer behaves consistently, that is, if he chooses bundle A in a situation in which bundle B was also available to him, he will not choose B in any other situation in which A is also available.

Symbolically,

If A > B, then B ≯ A

> = Greater than

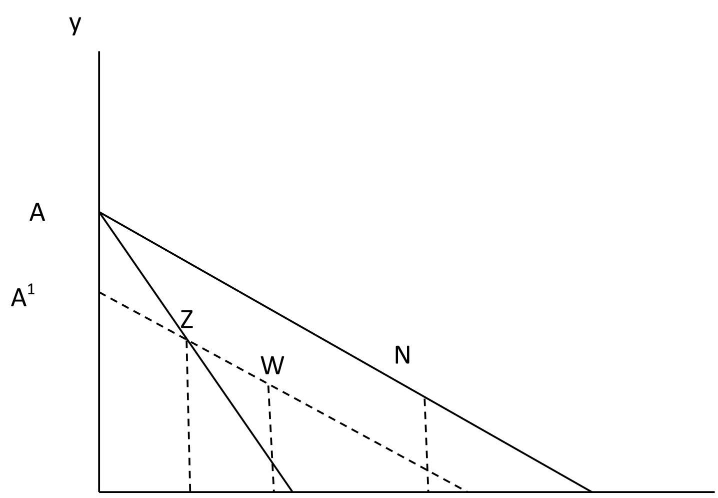
≯ = Not Greater than.

**Transitivity:** If any particular situation A > B and B > C, then A > C.

## The Revealed Preference Axiom

The consumer, by choosing a collection of goods in any one budget situation, reveals his preference for that particular collection. The chosen bundle is revealed preferred among all other alternative bundles available under the budget constraint.

Consider a situation whereby a consumer is faced with a budget line AB as shown in fig. 2.9, and chooses the collection of goods denoted by point Z, thus revealing his preference for this batch. Suppose that the price of x falls so that the new budget line facing the consumer is AC. It can be shown that the new batch will include a larger quantity of x.



0 $x_1$ $x_2$ B $x_3$ $B^1$ C x

**Fig. 2.9: The Revealed Preference**

Finally, we make a compensating variation of the income, which consists in the reduction of income so that the consumer has just enough income to enable him to continue purchasing Z if he so wishes. The compensating variation is shown in fig. 2.9 by a parallel shift of the new budget line so that the compensated budget line $A^1B^1$ passes through Z. Since the collection of Z is still available to him, the consumer will not choose any bundle to the left of Z on the segment $A^1Z$, because his choice would be inconsistent, given the original situation, all the batches on $A^1Z$ were revealed inferior to Z. Hence, the consumer will either continue to buy Z or he chooses a segment on the batch $ZB^1$, such as W, which includes a larger quantity of x (namely $x_2$). Secondly, if we remove the reduction in income and allow the consumer to move to the new budget line AC, he will choose a batch (such as N) to the right of W (if commodity x is normal with a positive income effect). The new revealed equilibrium (N) includes a larger quantity of x (i.e. $x_3$) resulting from fall in its price.

**The Weak and Strong Axiom of Revealed Preference (WARP and SARP)**

WARP is one of the criteria which need to be satisfied in order to make sure that the consumer is consistent with his preferences. If a bundle of goods A is chosen over another bundle B when both are affordable, then the consumer reveals that he prefers A over B. WARP says that when preferences remain the same, there are no circumstances (budget set) where the consumer strictly prefers B over A. By choosing A over B when both bundles are affordable, the consumer reveals that his preferences are such that he will never choose B over A regardless of income and prices.

If $A > B$, then $B \ngtr A$

**Note**: ⊁ is referred to as “Not preferred”

SARP is equivalent to the weak axiom of revealed preferences, except that the consumer is not allowed to be indifferent between the two bundles that are compared. That is, if WARP concludes that A ⊁ B, SARP goes a step further and conclude that A > B.

**Note**: ≥ is referred to as “greater than or equal to”.

If A is directly revealed preferred to B, and B is directly revealed preferred to C, then we say A is indirectly revealed preferred to C.

**Self-Assessment Exercise**

Analyse the concept of consistency in the consumer choice

**b. Indifference Curves**

The Cardinal Utility Approach is basically criticised for its assumption that utility is a measurable entity. Ordinalists believe that measurement of subjective utility on an absolute scale is neither possible nor necessary. This skepticism about the measurement of utility forced them to evolve an alternative approach to analyse consumer behaviour. The culmination of this search is the indifference curve theory. This approach does not require that consumers be able to measure utility in any specific units of measurement but does assume that consumers are able to rank their preferences for combinations of goods. An indifference curve is a locus of points, particular combinations or bundles of goods which yield the same utility (level of satisfaction) to the consumers, so that he is indifferent as to the particular combination he consumes.

**Assumptions**

**Complete:** The completeness assumption states that the consumer has a definite order of preferences or rank ordering for every conceivable combination of commodities. The completeness assumption is sometimes called the trichotomy assumption for the following reason. Suppose there are two different combinations of goods, A and B; this axiom assumes that the consumer must state – "I prefer A to B" or "I prefer B to A" or "A and B" are equally preferred." The consumer is not in a state of indecision. According to Hal R. Varian, "one might imagine extreme situations involving life or death choices where ranking the alternatives might be difficult or even impossible but such choices are for the most part outside the domain of economic analysis."

**Consistency or transitivity:** Along with complete ordering, we expect that the consumer preferences are not self-contradictory or conflicting with each other. This is the assumption of transitivity (the technical name for consistency is transitive). This means that if A is preferred to B, and B is preferred to C, then A is preferred over C.

Symbolically, if A > B, and B > C, then A > C.

**Non-Satiation:** The third important assumption is that consumers always prefer more goods to less. It means that the consumer is not satiated at least not in all goods. More precisely, if bundle A contains at least as much of every commodity as bundle B and more of at least one commodity, then A is preferred to B.

**Substitutability:** The goods consumed by the consumers are substitutable.

**Optimality:** Finally, the optimality assumption means the consumer tries to maximise utility given his income and market prices.

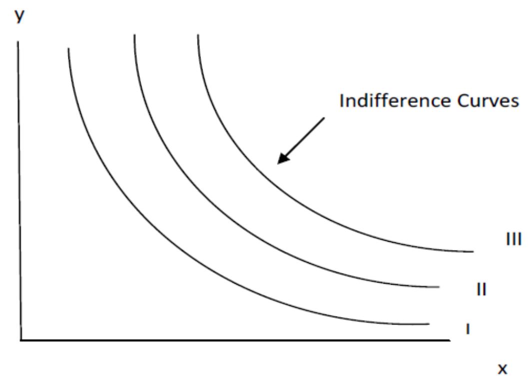



**Fig. 2.10a: Indifference Curve**

IC1 is a lower indifference curve, and thus offer lower level of satisfaction compare to ICII and ICIII. The higher the IC, the higher the level of satisfaction derived.

**Marginal Rate of Substitution (MRS)**

The shape of an indifference curve provides useful information about preferences. Indifference curve replaced the concept of marginal utility with the concept of marginal rate of substitution (MRS).

The negative slope of an indifference curve is called the marginal rate of substitution (MRS) of the two commodities concerned (say x and y), and is given by the slope of the tangent at that point.

$$slope\ of\ IC = -\frac{dy}{dx} = MRSx,y$$

According to Leftwich, “The marginal rate of substitution of X for Y ($MRS_{xy}$) is defined as the maximum amount of Y the consumer is willing to give up getting an additional unit of X and still remaining on the same indifference curve.” Thus, the marginal rate of substitution is just a fancy name for an acceptable trade-off between two goods.

**Diminishing Marginal Rate of Substitution (DMRS)**

Table 2.2 can be used to explain the concept of diminishing marginal rate of substitution. Since Y decrease and X increases, the change in Y is negative (i.e., -ΔY), so the equation is:

$MRS_{xy} = -\Delta Y/\Delta X$ and $MRS_{yx} = -\Delta X/\Delta Y$

However, as with price elasticity of demand, the convention is to ignore the minus sign
i.e., $MRS_{xy} = \Delta Y/\Delta X$ and $MRS_{yx} = \Delta X/\Delta Y$

In Table 2.2, oranges are denoted by X and apples by Y, we have the following marginal rate of substitution $MRS_x$ for $_y$

Between A and B: AA1/A1B = 6/3 = 2.0
Between B and C: BB1/B1C = 3/2 = 1.5
Between C and D: CC1/C1D = 4/10 = 0.4

Thus, $MRS_x$ for $_y$ diminishes with increasing X. This is the principle of diminishing marginal rate of substitution.

**Table 2.2: Diminishing Marginal Rate of Substitution**

| Combinations | X (Oranges) | Y (Apples) |
|---|---|---|
| A | 2 | 15 |
| B | 5 | 9 |
| C | 7 | 6 |
| D | 17 | 2 |

## Income and Substitution Effect of Price Change

The Slutsky's theorem states that "the substitution effect of a price change is always negative". A fall in the price of x from say p1 to p2 will lead to an increase in quantity demand from say x1 to x2. This effect is referred to as the total price effect which may be split into income effect and substitution effect. The substitution effect of a price change is the increase in the quantity bought as the price of the commodity falls, after adjusting income so as to keep the real purchasing power of the consumer the same as before. This adjustment in income is called compensating variation, and is shown graphically by a parallel shift of the new budget line until it becomes tangent to the initial indifference curve.

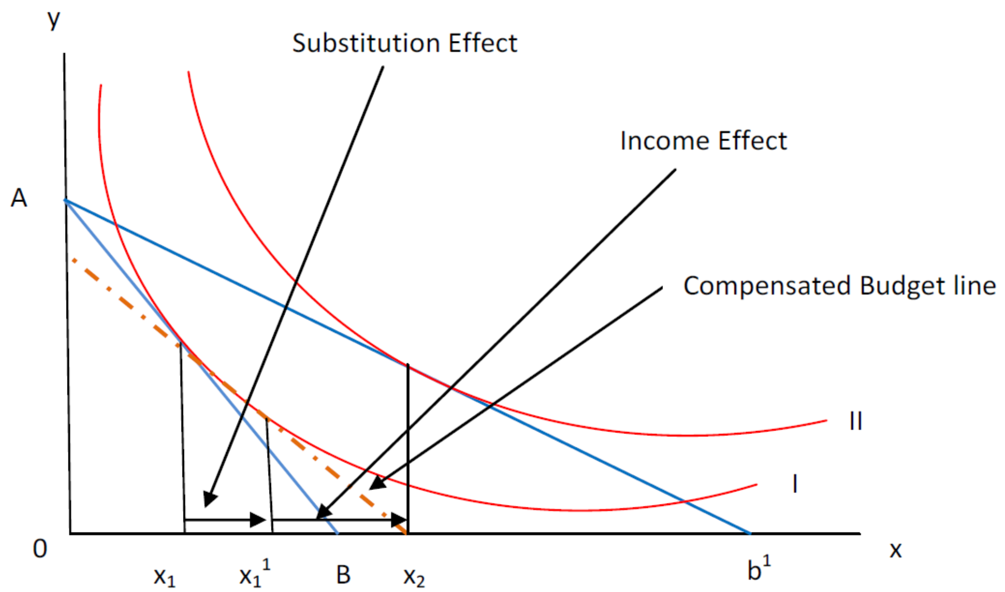



**Fig. 2.10b: Indifference Curve showing Income and Substitution Effect**

The purpose of the compensating variation is to allow the consumer to remain on the same level of satisfaction as before the price change. The compensated budget line will be tangent to the original indifference curve 1 at the first stage of equilibrium, because this line is parallel to the new budget line which is less steep than the original one when the price of x falls. Movement from $x_1$ to $x_1^1$ shows the substitution effect of the price change; the consumer buys more of x now that it is cheaper, substituting y for x.

The movement to a higher indifference curve was caused by a fall in the price of x, and the consumer now has a higher purchasing power, and if the commodity is normal, he will spend some of the increased real income on x, thus moving from $x_1^1$ to $x_2$. This is the income effect of the price change. The income effect of a price change is negative for normal goods and it reinforces the substitution effect of the diagram above. If, however, the commodity is inferior, the income effect of the price change will be positive; as the purchasing power increases, less of x will be bought.

## Self-Assessment Exercise

Analyse the indifference curve theory.

## 4.3.3 Diamond-Water Paradox

The diamond-water paradox poses the following perplexing observations: Even though water is obviously important to human activity (life cannot exist without water), the price of water is relatively low. Alternatively, diamonds are clearly much less important to human existence, but the price of diamonds is substantially higher. In other words, the utility obtained from water is obviously very great, while the utility obtained from diamonds is substantially less. The key question that arises is: Why are diamonds so much more expensive than water?

### Total Utility and Marginal Utility

Let us take a critical look into and clarification of the diamond-water paradox results by differentiating between total utility and marginal utility.

**Total Utility:** This is the overall satisfaction of wants and needs obtained from consuming a good. That is, total utility is the accumulated amount of satisfaction, or the total value, generated by several units of a good.

**Marginal Utility:** This is the extra satisfaction of wants and needs obtained from consuming one additional unit of a good. That is, marginal utility is the incremental satisfaction generated by, and the value of, a single unit of a good.

Water provides humans with an enormous amount of total utility. Water satisfies a lot of wants and needs for a lot of people. Water provides a high level of total utility because it is plentiful. However, because it is so plentiful, the marginal utility of water is relatively low. An extra liter of water provides very little additional satisfaction. In contrast, the total utility generated by diamonds is relatively limited. Diamonds do not provide much overall satisfaction of wants and needs, compared to water. Many humans spend their entire lives without achieving any satisfaction from diamonds. Diamonds have very little total utility because they are not nearly as plentiful as water. Most houses do not have hot and cold running diamonds. Most people do not drink eight glasses of diamonds a day, take showers in diamonds, or fill their Olympic-sized swimming pools with hundreds of gallons of diamonds. However, because they are less plentiful, the marginal utility of diamonds is relatively high. An extra ounce of diamonds provides a great deal of extra satisfaction.

**Solving the Paradox**

The apparent contradiction between price and utility is cleared up by distinguishing between marginal utility and total utility, and with the understanding that marginal utility and not total utility is the key to determining price. Moreover, this paradox can be turned on its head by considering what might happen should the relative abundance of water and diamonds change.

1. If water were as limited as diamonds, then the marginal utility and thus price would also be quite high. In fact, if water and diamonds were equally limited in supply, the price of water would likely be several times the price of diamonds.
2. If diamonds were as plentiful as water, then the marginal utility and price would also be quite low. If water and diamonds were equally abundant in supply, then the price of diamonds would likely be only a fraction of the price of water.

**Self-Assessment Exercise**
Why is diamond valued more than water?

## 4.3.4 Principle of Optimization

### a. Introduction

Decision-makers (e.g. consumers, firms, governments) in standard economic theory are assumed to be "rational". That is, each decision- maker is assumed to have a preference ordering over the outcomes to which her actions lead and to choose an action, among those feasible, that is most preferred according to this ordering. We usually make assumptions that guarantee that a decision-makers preference ordering is represented by a payoff function (sometimes called utility function), so that we can present the decision-makers problem as one of choosing an action, among those feasible, that maximises the value of this function. Whenever we try to explain the behaviour of human beings, we need to have a framework on which our analysis can be based. In economics, we use a framework built on the following two simple principles.

**The optimisation principle**: People try to choose the best patterns of consumption that they can afford.

**The equilibrium principle**: Prices adjust until the amount that people demand of something is equal to the amount that is supplied.

Let us consider these two principles. If people are free to choose their actions, it is reasonable to assume that they will try to choose things they want rather than things they don't want. Of course, there are exceptions to this general principle, but they typically lie outside the domain of economic behaviour. The second notion is a bit more problematic. It is at least conceivable that at any given time peoples' demands and supplies are not compatible, and hence something must be changing. These changes may take a long time to work themselves out, and, even worse, they may induce other changes that might "destabilise" the whole system.

This kind of thing can happen, but it usually doesn't. In the case of apartments, we typically see a fairly stable rental price from month to month. It is this equilibrium price that we are interested in, not in how the market gets to this equilibrium or how it might change over long periods of time. It is worth observing that the definition used for equilibrium may be different in different models. In the case of the simple market, we will examine here, the demand and supply equilibrium idea will be adequate for our needs. Typically, equilibrium will require that the economic agents' actions must be consistent with each other. How do we use these two principles to determine the answers to the questions we raised above?

**b. Demand Side Analysis**

Suppose that we consider all of the possible renters of the apartments and ask each of them the maximum amount that he or she would be willing to pay to rent one of the apartments. Assuming there is someone at the top who is willing to pay the highest price. Perhaps this person has a lot of money; perhaps he is very lazy and doesn't want to walk far. Suppose that this person is willing to pay ₦500 a month for an apartment. If there is only one person who is willing to pay ₦500 a month to rent an apartment, then if the price for apartments were ₦500 a month, exactly one apartment would be rented to the one person who was willing to pay that price.

Suppose that the next highest price that anyone is willing to pay is ₦490. Then if the market price were ₦499, there would still be only one apartment rented: the person who was willing to pay ₦ 500 would rent an apartment, but the person who was willing to pay ₦ 490 would not. And so, it goes. Only one apartment would be rented if the price were ₦498, ₦497, ₦496, and so on until we reach a price of ₦490. At that price, exactly two apartments would be rented: one to the ₦500 person and one to the ₦490 person. Similarly, two apartments would be rented until we reach the maximum price that the person with the third highest price would be willing to pay, and so on.

Economists call a person's maximum willingness to pay for something as that person's reservation price. The reservation price is the highest price that a given person will accept and still purchase the good. In other words, a person's reservation price is the price at which he or she is just indifferent between purchasing and not purchasing the good. In our example, if a person has a reservation price p it means that he or she would be just indifferent between living in the inner ring and paying a price p and living in the outer ring. Thus, the number of apartments that will be rented at a given price p* will just be the number of people who have a reservation price greater than or equal to p*. For if the market price is p*, then everyone who is willing to pay at least p* for an apartment will want an apartment in the inner ring, and everyone who is not willing to pay p* will choose to live in the outer ring.

Here, the price is depicted on the vertical axis and the number of people who are willing to pay that price or more is depicted on the horizontal axis. Another way to view the diagram is to think of it as measuring how many people would want to rent apartments at any particular price. Such a curve is an example of a demand curve-a curve that relates the quantity demanded to price. When the market price is above ₦500, zero apartments will be rented. When the price is between ₦500 and ₦490, one apartment will be rented. When it is between ₦490 and the third highest reservation price, two apartments will be rented, and so on. The demand curve describes the quantity demanded at each of the possible prices.

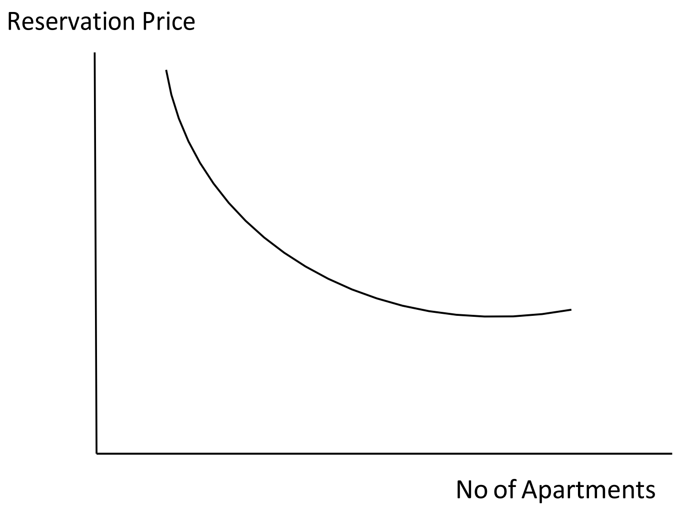



**Fig. 2.10c: Demand Side Analysis**

## Self-Assessment Exercise

Analyse the demand side optimization.

### c. Supply Side Analysis

Here we have to think about the nature of the market we are examining. The case to consider here is where there are many independent landlords who are out to rent their apartments for the highest price the market will bear. We will refer to this as the case of a competitive market. Other sorts of market arrangements are certainly possible. Consider the case where there are many landlords who all operate independently. It is clear that if all landlords are trying to do the best they can and if the renters are fully informed about the prices the landlords charge, then the equilibrium price of all apartments in the inner ring must be the same.

The argument is not difficult. Suppose instead that there is some high price, ph, and some low price, pl, being charged for apartments. The people who are renting their apartments for a high price could go to a landlord renting for a low price and offer to pay a rent somewhere between ph and pl. A transaction at such a price would make both the renter and the landlord better off. To the extent that all parties are seeking to further their own interests and are aware of the alternative prices being charged, a situation with different prices being charged for the same good cannot persist in equilibrium. But what will this single equilibrium price be? Consider a situation where a price is picked, and question on how many apartments will be supplied at that price was asked.

The answer depends to some degree on the time frame in which we are examining the market. If we are considering a time frame of several years, so that new construction can take place, the number of apartments will certainly respond to the price that is charged. But in the "short run"- within a given year, say the number of apartments is more or less fixed.

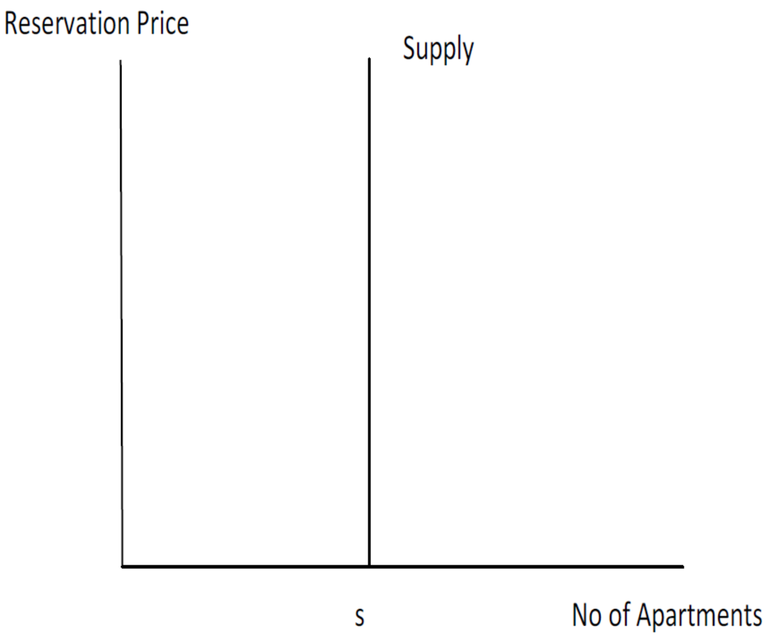



**Fig. 2.10d: Supply Side Analysis**

If we consider only this short-run case, the supply of apartments will be constant at some predetermined level. The supply curve in this market is depicted in fig. 2.10d as a vertical line. Whatever price is being charged, the same number of apartments will be rented, namely, all the apartments that are available at that time.

## Self-Assessment Exercise

Explain supply side optimization.

## d. Market Equilibrium

We now have a way of representing the demand and the supply side of the apartment market. Let us put them together and ask what the equilibrium behaviour of the market is. We do this by drawing both the demand and the supply curve on the same graph (Figure 2.10e).

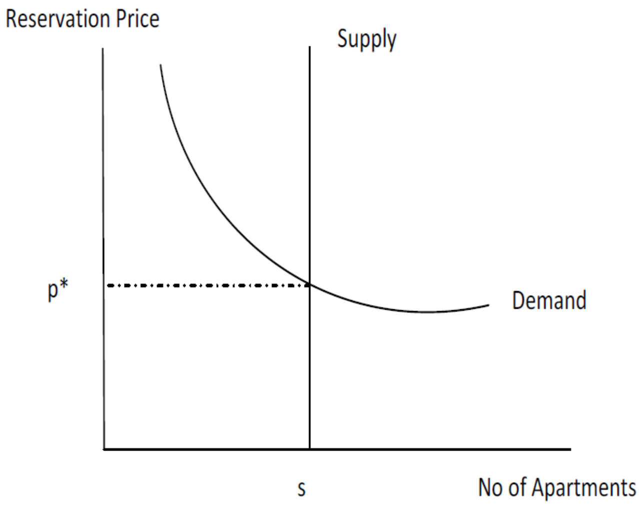



**Fig. 2.10e: Market Equilibrium**

In this graph (figure 2.10e), we have used p* to denote the price where the quantity of apartments demanded equals the quantity supplied. This is the equilibrium price of apartments. At this price, each consumer who is willing to pay at least p* is able to find an apartment to rent, and each landlord will be able to rent apartments at the going market price. Neither the consumers nor the landlords have any reason to change their behaviour. This is why we refer to this as equilibrium; no changes in behaviour will be observed.

To better understand this point, let us consider what would happen at a price other than p*. For example, consider some price $p < p^*$ where demand is greater than supply. Can this price persist? At this price, at least some of the landlords will have more renters than they can handle. There will be lines of people hoping to get an apartment at that price; there are more people who are willing to pay the price p than there are apartments. Certainly, some of the landlords would find it in their interest to raise the price of the apartments they are offering.

Similarly, suppose that the price of apartments is some p greater than p*. Then some of the apartments will be vacant: there are fewer people whoare willing to pay p than there are apartments. Some of the landlords are now in danger of getting no rent at all for their apartments. Thus, they will have an incentive to lower their price in order to attract more renters.

If the price is above p* there are too few renters; if it is below p* there are too many renters. Only at the price p* is the number of people who are willing to rent at that price equal to the number of apartments available for rent. Only at that price does demand equal supply. At the price p* the landlords and the renters' behaviours are compatible in the sense that the number of apartments demanded by the renters at p* is equal to the number of apartments supplied by the landlords. This is the equilibrium price in the market for apartments.

Once we have determined the market price for the inner-ring apartments, we can ask who ends up getting these apartments and who is exiled to the farther-away apartments. In the market equilibrium, everyone who is willing to pay p* or more gets an apartment in the inner ring, and everyone who is willing to pay less than p* gets one in the outer ring. The person who has a reservation price of p* is just indifferent between taking an apartment in the inner ring and taking one in the outer ring. The other people in the inner ring are getting their apartments at less than the maximum they would be willing to pay for them. Thus, the assignment of apartments to renters is determined by how much they are willing to pay.

**Self-Assessment Exercise**
How can equilibrium be reached in the apartment demand and supply scenario?

## 4.4 Conclusion

Based on the analysis so far, we can conclude that

First, according to the cardinal theory, utility can be quantified in terms of the money a consumer is willing to pay for it, thus $MU_x = P_x$.

Second, Ordinal theory states that utility cannot be measured in terms of monetary value due to changes in the utility of money.

Third, Revealed preference hypothesis states that if Good A is revealed preferred to Good B when Good A is available, then it will not happen that Good B is preferred to Good A when Good A is available. This is termed the principle of consistency or transitivity.

Fourth, higher indifference curve gives a higher satisfaction.

Fifth, changes in the price of good x will lead to a change in its quantity demanded, thus substituting between good x and y.

Finally, Goods x and y can be consumed in different orders that maximise the consumer's utility. Water is essential to human existence, yet commands low price compared to diamond.

## 4.5 Summary

This unit focused on consumer behaviour, and topics that capture how rational consumers behave were discussed. In particular, ordinal and cardinal utility theory was discussed, with special attention to the quantification of utility through marginal and diminishing marginal utility, before countering it with revealed preference hypothesis and the indifference curves. The concept of diamond-water paradox was visited in order to better understand the theory of value, and why diamond commands higher price than water which is essential to life. Finally, the principle of optimisation which focuses on the rationality of consumers and decision makers in the economy was discussed. Consumers want to maximise their utility while firms want to maximize their profits or minimise their cost of production. In order to analyse these two concepts, the demand and supply side analysis was applied to apartment demand in order to arrive at equilibrium.

## 4.6 Tutor-Marked Assignment

1. Suppose that there were 25 people who had a reservation price of ₦500, and the 26th person had a reservation price of ₦200. What would the demand curve look like?
2. If people have different reservation prices, why does the market demand curve slope down?
3. If the price of apartments increased in the short run from ₦500 to ₦520 under the supply side analysis, what will happen to the number of apartments available for rent?
4. Analyse the apartment market equilibrium.
5. If our model of rent control allowed for unrestricted subletting, who would end up getting apartments in the inner circle?

## 4.7 Reference/Further Reading

Koutsoyiannis, A. (1987.). *Modern Microeconomics*. (2nd ed.).

Varian, H . R. (2000.). *Intermediate Microeconomics: A Modern Approach*. (7th ed.).

Robert, S. P. & Rubinfeld, D. L. (2009). *Microeconomics.* (7th ed.). New Jersey: Pearson Education International.

Taylor, J. & Akila, W. (2012). Principles *of Economics*. (7th ed.). Mason, OH: South-Western Cengage Learning.

### 4.8 Answers to Self-Assessment Exercises

**Question**

Explain briefly what you understand by the term diminishing marginal utility.

**Model Answer**

Diminishing marginal utility implies that the utility derived from successive consumption of a unit of a commodity by a consumer diminishes. In other words, when an individual consumes more of a commodity, the total utility increases at a decreasing rate. However, after certain stage, the total utility also starts decreasing and the marginal utility becomes negative. This means that the individual does not need the commodity further. To put it simply, even the most beautiful place of the world or the sweetest music can make you feel bored after certain stage.

**Question**

Analyse the concept of consistency in the consumer choice

**Model Answer**

The consumer behaves consistently, that is, if he chooses bundle A in a situation in which bundle B was also available to him, he will not choose B in any other situation in which A is also available.

Symbolically,

If A > B, then B ≯A

> = Greater than

≯ = Not Greater than.

**Question**

Why is diamond valued more than water?

**Model Answer**

Water provides humans with an enormous amount of total utility. Water satisfies a lot of wants and needs for a lot of people. Water provides a high level of total utility because it is plentiful. However, because it is so plentiful, the marginal utility of water is relatively low. An extra liter of water provides very little additional satisfaction. In contrast, the total utility generated by diamonds is relatively limited. Diamonds do not provide much overall satisfaction of wants and needs, compared to water. Many humans spend their entire lives without achieving any satisfaction from diamonds. Diamonds have very little total utility because they are not nearly as plentiful as water. Most houses do not have hot and cold running diamonds. Most people do not drink eight glasses of diamonds a day, take showers in diamonds, or fill their Olympic-sized swimming pools with hundreds of gallons of diamonds. However, because they are less plentiful, the marginal utility of diamonds is relatively high. An extra ounce of diamonds provides a great deal of extra satisfaction.

**Question**

How can equilibrium be reached in the apartment demand and supply scenario?

**Model Answer**

**Market Equilibrium**

The demand and the supply side of the apartment market can be represented on the graph below to illustrate the equilibrium behaviour of the market.

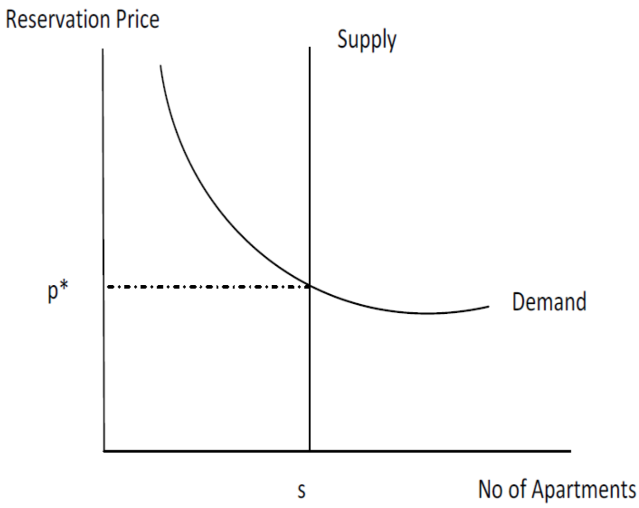



**Fig. 2.10e: Market Equilibrium**

In this graph (figure 2.10e), we have used p* to denote the price where the quantity of apartments demanded equals the quantity supplied. This is the equilibrium price of apartments. At this price, each consumer who is willing to pay at least p* is able to find an apartment to rent, and each landlord will be able to rent apartments at the going market price. Neither the consumers nor the landlords have any reason to change their behaviour. This is why we refer to this as equilibrium; no changes in behaviour will be observed.

To better understand this point, let us consider what would happen at a price other than p*. For example, consider some price p < p* where demand is greater than supply. Can this price persist? At this price, at least some of the landlords will have more renters than they can handle. There will be lines of people hoping to get an apartment at that price; there are more people who are willing to pay the price p than there are apartments. Certainly, some of the landlords would find it in their interest to raise the price of the apartments they are offering.

Similarly, suppose that the price of apartments is some p greater than p*. Then some of the apartments will be vacant: there are fewer people whoare willing to pay p than there are apartments. Some of the landlords are now in danger of getting no rent at all for their apartments. Thus, they will have an incentive to lower their price in order to attract more renters.

If the price is above p* there are too few renters; if it is below p* there are too many renters. Only at the price p* is the number of people who are willing to rent at that price equal to the number of apartments available for rent. Only at that price does demand equal supply. At the price p* the landlords and the renters' behaviours are compatible in the sense that the number of apartments demanded by the renters at p* is equal to the number of apartments supplied by the landlords. This is the equilibrium price in the market for apartments.

Once we have determined the market price for the inner-ring apartments, we can ask who ends up getting these apartments and who is exiled to the farther-away apartments. In the market equilibrium, everyone who is willing to pay p* or more gets an apartment in the inner ring, and everyone who is willing to pay less than p* gets one in the outer ring. The person who has a reservation price of p* is just indifferent between taking an apartment in the inner ring and taking one in the outer ring. The other people in the inner ring are getting their apartments at less than the maximum they would be willing to pay for them. Thus, the assignment of apartments to renters is determined by how much they are willing to pay.

# Module 3 Theory of Production and Cost



## Unit 1 Theory of Production

**Contents**



## 1.1 Introduction

Whether it is a big or small firm that markets a single product, suppliers usually face a difficult task. Producing an economic good or service requires a combination of land, labour, capital, and entrepreneurs. The theory of production deals with the relationship between the factors of production and the output of goods and services. The theory is generally based on the short run; a period of production that allows producers to change only the amount of the variable input called labour. This contrasts with the long run, a period of production long enough for producers to adjust the quantities of all their resources, including capital. For example, Ford

Motors hiring 300 extra workers for one of its plants is a short-run adjustment. If Ford builds a new factory, this is a long-run adjustment.

Put differently, the technical conditions facing the firm are summarised in the production function. The production function is a mathematical statement of the physical relationship between a firm's input of productive resources and its output of goods and services per unit of time, given available technology. In other words, a firm's production function establishes the relationship between the rate of flow of output and the rate of flow of corresponding inputs needed to produce it, given existing technology.

A production function can be expressed mathematically as:

$Q = f(X_1, X_2, X_3, \ldots, X_n)$

Where $X_1$, $X_2$, $X_3$, …,$X_n$ represents quantities of various inputs and Q represents the quantity of goods they can generate per unit of time.

## 1.2 Objectives

At the end of this unit, you should be able to:

- explain the theory of production
- describe the three stages of production
- explain the law of variable proportion
- explain the economics of scale.

## 1.3 Main Content

### 1.3.1 Law of Variable Proportions

The **law of variable proportions** states that, "in the short run, output will change as one input is varied while the others are held constant." For example, if you are preparing a meal, you know that a little bit of salt will make the food taste better. A bit more may make it tastier still. Yet, at some point, too much salt will ruin the meal. As the amount of the input (salt) varies, so does the output - the quality of the meal. The law of variable proportions deals with the relationship between the input of productive resources and the output of final products. This law helps answer the question: How is the output of the final product affected as more units of one variable input or resource are added to a fixed amount of other resources?

A farmer, for example, may have all the land, machines, workers, and other items needed to produce a crop. However, the farmer may have some questions about the use of fertilizer. How will the crop yield be affected if different amounts of fertilizer are added to fixed amounts of the other inputs? In this case, the variable input is the fertilizer added per acre. Of course, it is possible to vary all the inputs at the same time. The farmer may want to know what will happen to output if the fertilizer and other factors of production are varied. Economists do not like to do this, however, because when more than one factor of production is varied, it becomes harder to gauge the impact of a single variable on total output.

**Self-Assessment Exercise**

Explain the law of variable proportions.

### 1.3.2 Production Function

The law **of variable proportions** can be illustrated by using a production function **-** a concept that describes the relationship between changes in output to different amounts of a single input while other inputs are held constant. The production function can be illustrated with a schedule, such as Table 3.1. The production schedule in Figure 3.1 lists hypothetical output as the number of workers is varied from zero to 12. With no workers, for example, there is no output. If the number of workers increases by one, output rises to seven. Add yet another worker and total output rises to 20. This information can be used to construct the production function. In this example only the number of workers changed. No changes occur in the amount of machinery used, the level of technology, or the quantities of raw materials—unprocessed natural products used in production. Under these conditions, any change in output must be the result of the variation in the number of workers.

## Total Product

The second column in the production schedule in Table 3.1 shows total product, or total output produced by the firm. The numbers indicate that the plant barely operates when it has only one or two workers. As a result, some resources stand idle much of the time. As more workers are added, however, total product rises. More workers can operate more machinery, and plant output rises. Additional workers also mean that the workers can specialise. For example, one group runs the machines, another handles maintenance, and a third group assembles the products. By working in this way - as a coordinated whole - the firm can be more productive.

As more workers are added output continues to rise, but it does so at a slower rate until it can grow no further. Finally, the addition of the eleventh and twelfth workers cause total output to decline because these workers just get in the way of the others. Although the ideal number of workers cannot be determined until costs are considered, it is clear that the eleventh and twelfth workers will not be hired.

**Table 3.1: The Production Schedule**

| No. of Workers | Total Product | Average Product | Marginal Product | Stages of Production |
|---|---|---|---|---|
| 0 | 0 | 0 | 0 | |
| 1 | 7 | 7 | 7 | |
| 2 | 20 | 10 | 13 | Stage I |
| 3 | 38 | 12.6 | 18 | |
| 4 | 62 | 15.5 | 24 | |
| 5 | 90 | 18 | 28 | |
| | | | | |
| 6 | 110 | 18.3 | 20 | |
| 7 | 129 | 18.3 | 19 | |
| 8 | 138 | 17.3 | 9 | Stage II |
| 9 | 144 | 16 | 6 | |
| 10 | 148 | 14.8 | 4 | |
| | | | | |
| 11 | 145 | 13.2 | -3 | Stage III |
| 12 | 135 | 11.3 | -10 | |

**Average Product**

The quantity of total output produced per unit of a variable input, holding all other inputs fixed. Average product (AP) is found by dividing total product by the quantity of the variable input. Average product, which is also known as Average Physical Product (APP), is one of two measures derived from total product. Put differently, average product is the per unit production of a firm. Conceptually, it is simply the arithmetic mean of total product calculated for each variable input over a whole range of variable input quantities. Average product is generally considered less important than total product and marginal product in the analysis of short-run production.

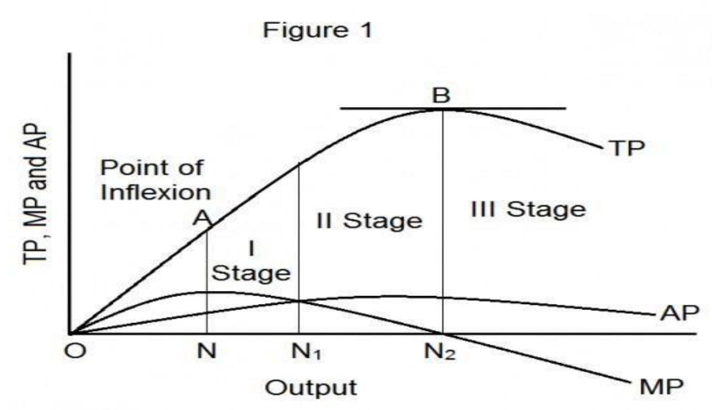



**Fig. 3.1: Production Function**

**Marginal Product**

The measure shown in the third column of the production schedule above is an important concept in economics. The measure is known as marginal product, the extra output or change in total product caused by the addition of one more unit of variable input. As we can see in the Figure 3.1, the marginal product, or extra output of the first worker, is seven. Likewise, the marginal product of the second worker-which is equal to the change in total product - is 13. Together, both workers account for 20 units of total product. When it comes to determining the optimal number of variable units to be used in production, changes in marginal product are of special interest. The production schedule shows the three stages of production - increasing returns, diminishing returns, and negative returns - that are based on the way marginal product changes as the variable input of labour is changed.

In Stage I, the first workers hired cannot work efficiently because there are too many resources per worker. As the number of workers increases, they make better use of their machinery and resources. This results in increasing returns (or increasing marginal products) for the first five workers hired. As long as each new worker hired contributes more to total output than the worker before, total output rises at an increasingly faster rate. Because marginal output increases by a larger amount every time a new worker is added, Stage I is known as the stage of increasing returns. Companies, however, do not knowingly produce in Stage I for very long. As soon as a firm discovers that each new worker adds more output than the last, the firm is tempted to hire another worker.

In Stage II of production the total production keeps growing but by smaller and smaller amounts. Any additional workers hired may stock shelves, package parts, and do other jobs that leave the machine operators free to do their jobs. The rate of increase in total production, however, is now starting to slow down. Each additional worker, then, is making a diminishing, but still positive, contribution to total output. Stage II illustrates the principle of diminishing returns, the stage where output increases at a diminishing rate as more units of a variable input are added. In the above table, Stage II begins when the sixth worker is hired, because the 20-unit marginal product of that worker is less than the 28-unit marginal product of the fifth worker.

The third stage of production begins when the eleventh worker is added. By this time, the firm has hired too many workers, and they are starting to get in each other's way. Marginal product becomes negative and total plant output decreases. Most companies do not hire workers whose addition would cause total production to decrease. Therefore, the number of workers hired would be found only in Stage II. The exact number of workers hired depends on the cost of each worker. If the cost is low, the firm should hire at least six, but no more than 10, workers.

**Self-Assessment Exercise**

What is the difference between total product, average product and marginal product?

### 1.3.3 Technological Progress

As knowledge of new and more efficient methods of production become available, technology changes. Furthermore, new inventions may result in the increase of the efficiency of all methods of production. At the same time, some techniques may become inefficient and drop out from the production function. These changes in technology constitute technological progress. Graphically, the effect of innovation in processes is shown with an upward shift of the production function in Figure 3.2(a) or a downward movement of the production isoquants graph in Figure 3.2(b). This shift shows that the same output may be produced by fewer factor inputs, or more output may be obtained with the same inputs.

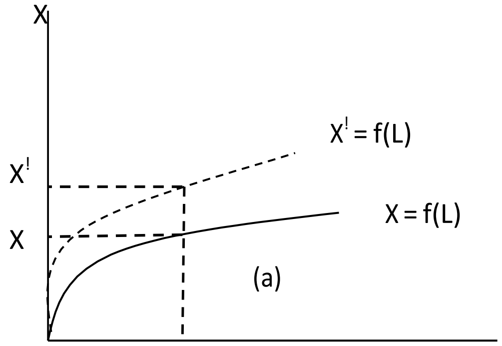



**Fig. 3.2a: Technological Progress**

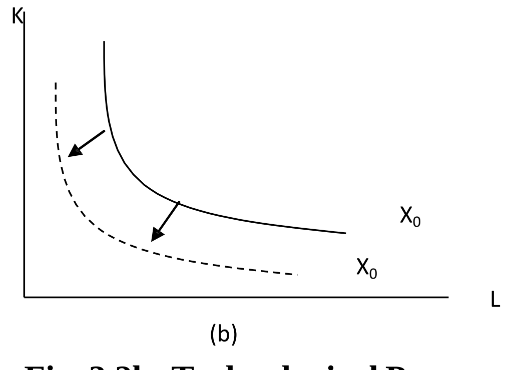



**Fig. 3.2b: Technological Progress**

Technical progress may also change the shape of the isoquants. Hicks have distinguished three types of technical progress, depending on its effect on the rate of substitution of the factors of production. They are:

1. Capital-deepening technical progress
2. Labour-deepening technical progress
3. Neutral-technical progress

**Self-Assessment Exercise**
What are the effects of technological progress?

### 1.3.4 Returns to Scale

There are three aspects of return to scale; increasing returns to scale, constant returns to scale and decreasing returns to scale.

**Increasing returns to scale**
This is a situation where a given proportional change in all resources in the long run results in a proportional greater change in production. Increasing returns to scale exists if a firm increases all resources, labour, capital, and other inputs by a given proportion (say 10 per cent) and output increases by more than this proportion (that is more than 10 per cent). If output is more than doubles when inputs are doubled, there are increasing returns to scale. This might arise because the larger scale of operation allows the managers and workers to specialise in their tasks and to make use of more sophisticated, large-scale factories and equipment. The prospect of increasing return to scale is an important issue from a public-policy perspective. If there are increasing returns, then it is economically advantageous to have one large firm producing (at relatively low cost) rather than to have many small firms produce.

**Constant Returns to Scale**
A second possibility with respect to the scale of production is that output may double when inputs are doubled. In this case, we say there are constant returns to scale. With constant returns to scale, the size of the firm’s operation does not affect the productivity of its factors: Because one plant using a particular production process can easily be replicated, two plants produce twice as much output. For example, a large travel agency might provide the same service per client and use the same ratio

of capital (office space) and labour (travel agents) as a small agency that services fewer clients.

**Decreasing Returns to Scale**

Finally, output may be less than double when all inputs double. This case of decreasing returns to scale applies to some firms with large-scale operations. Eventually, difficulties in organising and running a large- scale operation may lead to decreased productivity of both labour and capital. Communication between workers and managers can become difficult to monitor as the workplace becomes more impersonal. Thus, the decreasing returns case is likely to be associated with the problem of coordinating tasks and maintaining a useful line of communication between managers and workers.

**Self-Assessment Exercise**

List and explain the economies of scale.

### 1.3.5 Substitution among Inputs

With two inputs that can be varied, a manager will want to consider substituting one input for another. The slope of each isoquant indicates how the quantity of one input can be traded off against the quantity of the other, while output is held constant. When the negative sign is removed, we call the slope the *marginal rate of technical substitution* (*MRTS*). The marginal rate of technical substitution of labour for capital is the amount by which the input of capital can be reduced when one extra unit of labour is used, so that output remains constant. This is analogous to the marginal rate of substitution in consumer theory. The MRTS is always measured as a positive quantity.

MRTS = **-** Change in capital input ÷ Change in labour input

$$\Delta K/\Delta L \text{ (for a fixed level of q)}$$

Where $\Delta K$ and $\Delta L$ are small changes in capital and labour along an isoquant.

**Diminishing Marginal Rate of Technical Substitution**

With diminishing MRTS, that is, the MRTS falls as we move down along the isoquants. The mathematical implication is that isoquants, like indifference curves, are convex, or bowed inward. This is indeed the case for most production technologies. The diminishing MRTS tells us that the productivity of any one input is limited. As more and more units of labour are added to the production process in place of capital; the productivity of labour falls. Similarly, when more capital is added in place of labour, the productivity of capital falls. Production needs a balanced mix of both inputs.

Based on the above discussion, we can infer that the MRTS is closely related to the marginal products of labour $MP_L$ and capital $MP_K$. To see how, imagine adding some labour and reducing the amount of capital sufficient to keep output constant. The additional output resulting from the increased labour input is equal to the additional output per unit of additional labour times the number of units of additional labour.

Additional output from increased use of labour = $(MP_L)*(\Delta L)$

Similarly, the decrease in output resulting from the reduction in capital is the loss of output per unit reduction in capital times the number of units of capital reduction:

Reduction in output from decreased use of capital = $(MP_K)*(\Delta K)$

As we are keeping output constant by moving along an isoquant, the total change in output must be zero. Thus,

$(MP_L)*(\Delta L) + (MP_K)*(\Delta K) = 0$

Now, by rearranging terms, we see that: $(MP_L) / (MP_K) = - (\Delta K/\Delta L) = MRTS$

The above equation tells us that the marginal rate of technical substitution between two inputs is equal to the ratio of the marginal products of the inputs.

**Self-Assessment Exercise**

Differentiate between MRTS and MRS.

## 1.4 Conclusion

From our discussion so far on the theory of production, we can infer the following:

First, a change in the variable input called labour results in a change in production

Second, the law of variable proportions states that output will change as one input is varied while others are held constant

Third, as the number of labour units increases with capital held constant, the output increases, and decreases after some time

Fourth, as knowledge of new and more efficient method of production become available, technology changes

Fifth, if output more than doubles when inputs is doubled, there is increasing returns to scale. If otherwise, there is decreasing returns to scale

Finally, the slope of isoquants for two substitute inputs indicates how the quantity of one input can be traded off against the quantity of the other while output is held constant.

## 1.5 Summary

The law of variable proportions helped to establish that when one input is increased while others are held constant, output will increase in the short run. This law deals with the relationship between the inputs of productive resources and the final product. The production function describes the relationship between changes in output to different amounts of a single input while other inputs are held constant. The concept of total and marginal product and the effect of technological progress on production were discussed.

## 1.6 Tutor-Marked Assignment

1. State the theory of production and law of variable proportions.
2. Explain the meaning of the following:
   - production function
   - total product
   - marginal product
   - stages of production

- diminishing returns.

3. Explain how marginal product changes in each of the three stages of production.
4. Identify what point will eventually be reached if companies continue adding workers.
5. Provide an example of a time when you entered a period of diminishing returns or even negative returns. Explain why this might have occurred.

## 1.7 References/Further Reading

Gary, E. C. (2001). *Principles and Practices of Economics.* (Interactive student ed.). New York: McGraw-Hill.

Koutsoyiannis, A. (1987). *Modern Microeconomics*. (2nd ed.).

Robert, S. P. & Rubinfeld, D. L. (2009). *Microeconomics.* (7th ed.). New Jersey: Pearson Education International.

Tucker, I. B. (2014). *Economics for Today*. (8th ed.). Mason, OH: South-Western Cengage Learning.

## 1.8 Answers to Self-Assessment Exercises

**Question**

What are the effects of technological progress?

**Model Answer**

The effects of technological progress can be summarized as follow:

First, technological progress may result in the increase of the efficiency of all methods of production. This could manifest in the form of producing the same output with fewer factor inputs, or more output may be obtained with the same inputs.

Second, technological progress may also change the shape of the isoquants. This is in relation to cost of production, which could be less as technology can lead to reduction in the number and quantity of labour employed with greater efficiency. It is on this basis that Hicks distinguished three types of technical progress, depending on its effect on the rate of substitution of the factors of production. They are:

1. Capital-deepening technical progress
2. Labour-deepening technical progress
3. Neutral-technicalprogress

# Unit 2 Theory of Cost

**Contents**



## 2.1 Introduction

Economists generally emphasize costs to be one of the major elements that profit maximizing firms seek to control closely. The overall yardstick for the measurement of managerial efficiency in a firm is profitability, and profit is basically the difference between cost and revenue. To any firm, the basic factor underlying its willingness and ability to supply a commodity is its cost of production. Costs are very important elements in supply decisions since the production of any commodity involves the use of economic resources that are scarce and thus command a price, which constitutes costs to the firm. A firm that manages its costs prudently can succeed in reducing its unit cost over those of competing firms in the industry, as this offers the firm a competitive edge over others in the industry and greater opportunity for profit making.

## 2.2 Objectives

At the end of this unit, you should be able to:

- differentiate between implicit and explicit cost
- explain the difference between economic profit and accounting profit
- define the concept of marginal and average cost

- describe internal and external economies of scale.

## 2.3 Main Content

### 2.3.1 Implicit and Explicit Costs

**Explicit Costs**

Explicit costs comprise those clear-cut historical payments or money outlays for inputs bought or hired by the firm and are documented in the firm's expenses recorded during a given production period. Example of explicit costs includes wages and salaries paid, rent, interest expense and so on.

**Implicit Costs**

Implicit costs are the opportunity costs associated with self-employed resources in the business. They comprise all of the remuneration that owner-employed resources in the firm could have earned in alternative employment if the owners had not engaged them in their own business. Examples of implicit costs include the implicit rent that could have been paid in the case of business premises owned by the proprietor, implicit wages and salaries for the owner and his wife, for example, who have spent much of their time working in the business but without a stipulated remuneration and so on.

**Accounting vs. Economic Profit**

The distinction between accounting and economic profit stems from the difference between implicit and explicit costs. To the accountant, profit is simply the excess of total revenue over historical cost, i.e., cost of the firm for which actual payments have been effected in the past or for which the firm is committed to future payments. This accountant's restriction on costs to explicit cost has the advantage of ensuring objectivity.

Records of transaction exist, backed up with receipts for payment made, and appropriate documentation for credit transaction. A business profit defined as the difference between revenue and explicit costs is generally referred to as accounting profit. However, an Economist considers the concept of historical costs to be too narrow when estimating the total cost and the profits of running a business. Therefore, economists consider profits to be the excess of total revenue over the explicit and implicit opportunity costs of the firm. From the above discussion, we can infer that a firm considered profitable from the accountant's viewpoint may actually be suffering losses from the economist's viewpoint.

For example, examine the profit position of Ajibola and Sons Ltd., whose revenue and cost data are given as:

| | | | |
|---|---|---|---|
| Total revenue | ₦40, 000 | | |
| Less Cost of goods sold | ₦22, 000 | | |
| Gross Profit | …………………………………… | | ₦18, 000 |
| Less Operating Expenses: | | | |
| Fixed Cost | ₦8, 000 | | |
| Variable Costs | ₦2, 000 | ₦10, 000 | |
| Net Income (Accounting Profit) | …………………………. | | ₦8, 000 |
| Less: | | | |
| Implicit Rental Expense | ₦2, 000 | | |
| Implicit Salary Expense | ₦4, 000 | | |
| Implicit Interest Expense | ₦500 | ₦6, 500 | |
| Economic Profit | ……………………………………………… | | ₦1, 500 |

**Cost vs. Outlay**

All cash outlay or expenditure does not necessarily represent cost. Expenditure on items such as automobiles, air-conditioners, engine oil, meat pie, etc. cannot be classified as cost because some of them can last for a long time. For example, if a firm decides to purchase a new car for ₦200, 000, its wealth position or net worth remains unaltered. It has merely substituted a financial asset (cash) with a real asset (car). If 12 months later, it decides to sell the car for ₦180, 000, the cost of using the car for 12 months is ₦20, 000. If the firm chooses to keep the car until it is fully depreciated, then the cost of the car to the firm each period is the amount of depreciation apportioned to that price. Cash outlays for durable assets are not necessarily costs because there are still opportunities for some residual alternatives ahead. The car for example could be sold at the end of the second year for ₦160, 000 and the money can be put into alternative uses. In the case of consumable good like meat-pie, their service is fully utilised almost immediately and thus they are classified as costs.

A natural starting point in cost analysis is to assume the existence of a direct, well-defined relationship between the costs of production and the level of output. This relationship between cost and output at various levels is known as the cost function. Thus, the cost function of production depends on the level of output.

This can be expressed algebraically as:

$C = f(Q)$

Where C = naira cost of production

Q = quantity

Note: output rate depends on inputs. This can be expressed as:

$Q = f(I)$

Where I = resources/inputs.

The production function establishes the technical relationship between input and output flows. Thus, given input prices and the relevant production function, the cost of any giving level of output can be computed. We can therefore say precisely that the cost behaviour of a firm depends on the character of its production function and the prices of its inputs. And the nature of production function determines the shape of the total cost curve.

**Self-Assessment Exercise**

Differentiate between implicit and explicit cost, economic and accounting profit.

### 2.3.3 Marginal and Average Cost

**Marginal Cost**

Marginal cost is the change in total cost divided by the corresponding change in the output, or the rate of change of total cost as output changes. The change in total cost may be an increase or a decrease brought about by an increase or a decrease in output. Marginal cost (MC) is thus given or measured by:

$$MC = \frac{\Delta TC}{\Delta Q} \text{ or } \frac{d(TC)}{dQ}$$

Where $^{d}$TC = Change in total cost

DQ = Change in output, d(TC)/dQ = The rate of change of total cost as output changes. It is the derivative or slope of the total cost curve. Because of the fixed cost (FC) component in total cost (TC), there is economy of scale in the short run as output increases.

Note that total cost (TC) = Fixed Cost (FC) + Variable Cost (VC).

**Variable Costs**

Variable costs are the costs of such inputs as raw materials, intermediate inputs and labour, all of which vary as output varies over time.

**Marginal Cost**

Marginal cost is the change in total cost that comes from making or producing one additional item. The purpose of analysing marginal cost is to determine at what point an organisation can achieve economies of scale. The calculation is most often used among manufacturers as a means of isolating an optimum production level. Put differently, manufacturing concerns often examine the cost of adding one more unit to their production schedules. This is because at some point, the benefit of producing one additional unit and generating revenue from that item will bring the overall cost of producing the product line down. The key to optimising manufacturing costs is to find that point or level as quickly as possible.

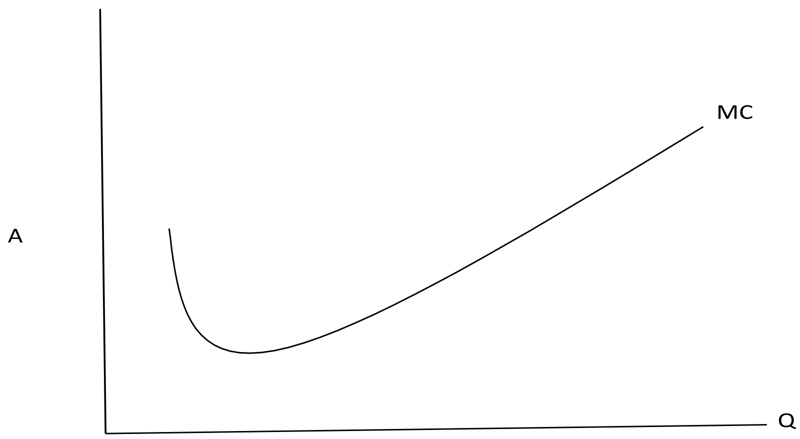



**Fig. 3.3a: Marginal Cost**

**Fixed Costs**

Fixed costs, indirect costs or overheads are business expenses that are not dependent on the level of goods or services produced by the business. They tend to be time-related, such as salaries or rents being paid per month, and are often referred to as overhead costs.

Fixed costs are not permanently fixed; they will change over time, but are fixed in relation to the quantity of production for the relevant period. For example, a company may have unexpected and unpredictable expenses unrelated to production; for instance, warehouse costs and the like are fixed only over the time period of the lease. There are no fixed costs in the long run, because the long run is a sufficient period of time for all short-run fixed inputs to become variable.

**Average Costs**

The average cost (AC) or cost per unit of output will first fall and eventually rise. It will initially fall because the fixed costs can be spread over larger and larger outputs, thus leading to lower and lower overhead cost per unit of output. However, average cost will rise because of the diminishing marginal productivity of variable factors which are being increased while fixed factors are kept constant. It may also be due to the rising prices and costs of the scarce inputs or raw materials as more and more are demanded. So, the AC curve will first fall from a high level and eventually rise, giving a U-shape. The relationship between marginal cost and average cost is such that when average cost is falling, marginal cost is below average cost, and when average costs is rising, marginal cost is above it. All this implies that marginal costs cut average costs at the minimum or turning point of average cost. In other words, when marginal cost is less than average cost, it pulls down average costs, and when marginal cost is higher than average cost, it pulls up average cost.

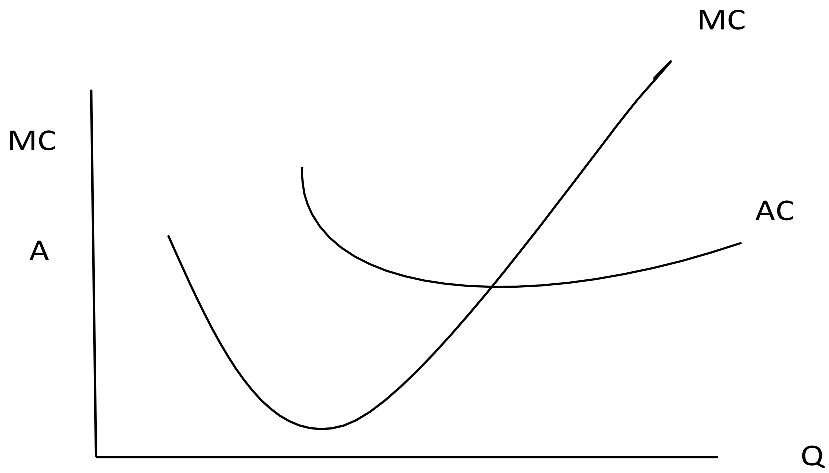



**Fig. 3.3b: Marginal and Average Cost**

Another aspect of average cost is its behaviour in the long run. In the short run, some costs are fixed relative to the installed production capacity. In the long run all costs are variable since the production capacity of plant size can be varied to suit demand. Thus, the long run average cost (LAC) will be the locus of points of the lowest attainable average cost of production for every output level using various plants sizes. Thus, LAC curve will envelop the short run average costs curves touching each of them underneath. This is illustrated in the diagram below.

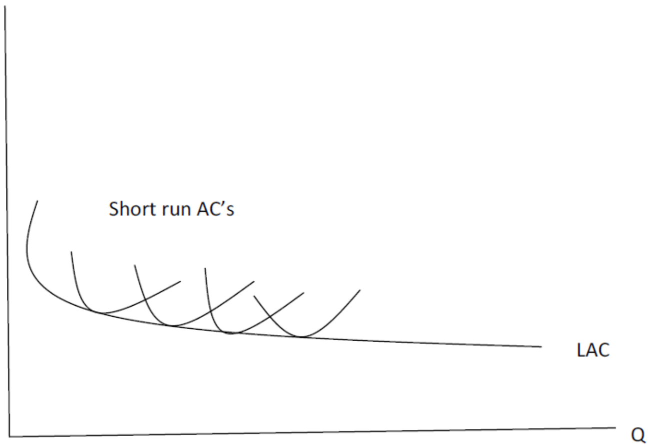



**Fig. 3.3c: Average Cost**

When average cost decreases with input using appropriate plant sizes, the industry is said to be characterised by decreasing cost or increasing returns and when average costs are constant, we have constant returns. But when long run average cost rises with output, we have increasing cost or decreasing returns in the industry. Since marginal cost is less than average cost when average cost is falling, it means that a situation when marginal cost is less than average cost implies decreasing costs or increasing returns or existing economy of scale.

**Self-Assessment Exercise**

Explain the concept of average cost?

## 2.3.4 Internal and External Economies of Scale

Economies or diseconomies of scale are said to be internal to the firm if the firm can cause them by enlarging its scale of production or plant size. Economies of scale are external if they result from the action of the entire industry and cannot be attained by a single firm changing its scale of production. Internal economies involve movements along given cost function of the firm. External economies entail a shift in the entire cost function. Internal and external economies or diseconomies of scale can otherwise be referred to as internal or external effects.

An external effect is said to exist when a firm's production function or cost function contains as a variable the output of the entire industry. Thus, an external effect is said to be experienced if the overall expansion of all firms in the industry alters the production of the individual in such a manner as to increase or reduce output or reduce the cost function. The simultaneous expansion of output by most firms in an industry may result in driving up the prices of certain key raw materials. This will entail extra cost for raw materials on all the firms. Similarly, if several firms are located in a geographical zone, this may attract the extension of railroad services to that region. This will cheapen the cost of transportation to all the firms in the area. This could not have occurred if just one or a few firms expanded their scale of production.

**Self-Assessment Exercise**

What is the effect of external economies of scale?

### 2.3.5 Constrained Cost Minimization

A manufacturing firm may wish to determine the minimum amount to spend on variable inputs to achieve a given level of output. In this regard, the function to be minimised is the cost function. The constraint equation is the production function. Let the cost function be given as:

$$C = f(Q) = P1X1 + P2X2 \qquad (1)$$

The production function is given as:

$$Q = f(X1, X2) \qquad (2)$$

The related Lagrangian function can be written as:

$$L = (P1X1 + P2X2) + \lambda[Q - f(X1, X2)] \qquad (3)$$

The first order condition requires that:

$$\frac{\partial L}{\partial X1} = P1 - \lambda f1 = 0 \qquad (4)$$

$$\frac{\partial L}{\partial X2} = P2 - \lambda f2 = 0 \qquad (5)$$

$$\frac{\partial L}{\partial \lambda} = Q - f(X1, X2) = 0 \qquad (6)$$

From equation (4) and (5),

$P1 = \lambda f1$

$P2 = \lambda f2$

$$\frac{P1}{P2} = \frac{f1}{f2} = \frac{MPx1}{MPx2}$$

Thus, the first order condition that the price ratio of the inputs must equal the ratio of their marginal products is met. This is to say that the slope of the iscost must equal the slope of the related isoquants.

## 2.4 Conclusion

From our discussion so far on the theory of firm, we can infer the following:

First, Cost is a major element of a profit maximising firm.

Second, Explicit cost comprises of historical outlays for inputs bought or hired by the firm which are recorded as expenses during a production period. Implicit costs are the opportunity costs associated with self-employed resources in the firm which could have earned alternative employment somewhere else.

Third, Expenditure on automobile, air conditioners, engine oil etc. is not to be classified as cost.

Fourth, Marginal cost is the change in total cost divided by the corresponding change in the output. Average cost will fall first and eventually rise because the fixed costs can be spread over larger outputs which lower the overhead cost per unit output.

Fifth, Constrained cost minimisation aims at the minimisation of cost in order to attain a given level of output.

## 2.5 Summary

The concept of implicit and explicit cost was discussed. The accounting and economic profit concept was reviewed. From an accounting point of view, once marginal revenue is greater than the cost, the company is making profit. However, from an economist's point of view, when marginal revenue is equal to marginal cost, the firm is in equilibrium, and thus competitive profit exists for the firm. Variable costs are those cost of inputs which may vary over time, while fixed costs are the costs of machinery and other fixed assets.

## 2.6 Tutor-Marked Assignment

1. Differentiate between accounting profit and economic profit
2. Analyse the concept of marginal and average cost
3. What do we mean by internal and external economies of scale?

## 2.7 References/Further Reading

Gary, E. C. (2001). *Principles and Practices of Economics*. (Interactive student ed.). New York: McGraw-Hill.

Miller, R. L. (2014). *Economics Today*. (17th ed.). Upper Saddle River, NJ: Pearson.

O'Sullivan, et *al*. (2014). Economics*: Principles, Applications, and Tools*. (8th ed.). Upper Saddle River, NJ: Pearson.

Parkin, M. (2014). Economics. (11th ed.). Upper Saddle River, NJ: Pearson.

Robert, S. P. & Daniel, L. R. (2009). *Microeconomics.* (7th ed.). New Jersey: Pearson Education International.

Sloman, J, (2006). *Economics* (6th ed,). Prentice Hall, London and New York

**2.8 Answers to Self-Assessment Exercises**

**Question**

What is the effect of external economies of scale?

**Model Answer**

Economies of scale are external if they result from the action of the entire industry and cannot be attained by a single firm changing its scale of production. External economies entail a shift in the entire cost function.

An external effect is said to exist when a firm's production function or cost function contains as a variable the output of the entire industry. Thus, an external effect is said to be experienced if the overall expansion of all firms in the industry alters the production of the individual in such a manner as to increase or reduce output or reduce the cost function. The simultaneous expansion of output by most firms in an industry may result in driving up the prices of certain key raw materials. This will entail extra cost for raw materials on all the firms. Similarly, if several firms are located in a geographical zone, this may attract the extension of railroad services to that region. This will cheapen the cost of transportation to all the firms in the area. This could not have occurred if just one or a few firms expanded their scale of production.

# Unit 3 Labour-Leisure Tradeoff

**Contents**



## 3.1 Introduction

Leisure is considered one good (often put on the horizontal-axis) and consumption is considered the other good. Since a consumer has a finite amount of time, he must make a choice between leisure (which earns no income for consumption) and labour (which does earn income for consumption).

The total amount of time that an individual has to allocate is known as his time endowment (denoted by *T*). The amount an individual allocates to labour (denoted by *L*) and leisure (*l*) is constrained by *T* such that:

$$l + L = T$$

or

$$l + (T - l) = T$$

A person's consumption is the amount of labour they choose multiplied by the amount they are paid per hour of labour (their wage, denoted by w). Thus, the amount that a person consumes is:

$$C = w(T - l)$$

When a consumer chooses no leisure

$(l = 0)$

Then $T - l = T$

and $C = wT$

From this labour-leisure tradeoff model, the substitution effect from various changes in price caused by welfare benefits, labour taxation, or tax credits can be analysed.

## 3.2 Objectives

At the end of this unit, you should be able to:

- explain the relationship between labour and leisure
- analyse the backward bending and inverted S-shaped labour supply curve
- state the assumptions and caveat.

## 3.3 Main Content

### 3.3.1 The Model

The labour supply decision is regarded as the ultimate trade-off between consumption and leisure. The advantage of working more is increased consumption; by working more hours, you earn more income and can purchase and consume more goods. The disadvantage of working more is decreased leisure; by working more hours, you have fewer hours to spend outside of work to sleep, raise children, watch TV, etc. (The economic definition of leisure is any time spent outside of work). Therefore, the goal of choosing how many hours you want to work-i.e., choosing your personal quantity of labour supplied-is to choose the consumption-leisure combination available to you that you like best.

For example, Shola can earn a wage of ₦10 an hour, and there’s 24 x 7 = 168 hours in the week. If she works 20 hours a week, she can consume 20 x ₦10 = ₦200 in goods and services a week and have 168 – 20 = 148 hours of leisure time left. However, if she works 40 hours a week, she can consume 40 x ₦10 = ₦400 in goods and services a week but only has 168 – 40 = 128 hours of leisure time left. This means that Shola's choice between working (a) 20 hours a week and (b) 40 hours a week is actually a choice between (a) consuming ₦200 in goods and services and enjoying 148 hours of leisure and (b) consuming ₦400 in goods and services and enjoying 128 hours of leisure. Which will Shola pick? She will pick the consumption-leisure combination that she likes the best. If she is the easygoing type, she probably likes leisure more than consumption, will pick (a), and work 20 hours a week. If she is the driven type, she probably likes consumption more than leisure, will pick (b), and work 40 hours a week.

Of course, Shola certainly has more choices than working 20 or 40 hours a week. She could work 30 hours a week, or 50, or not at all. Let's find out what Shola's consumption and leisure are per week at each number of hours Shola works per week (or, alternately, at each of Shola's personal quantity of labour transacted):

**Table 3.1: Shola's Consumption-Leisure Possibilities, No Non-labour Income ₦-Naira**

| Weekly Labour | Wage | Weekly Consumption | Weekly Leisure |
|---|---|---|---|
| | | = W*L | =168 hours - L |
| 0 | ₦10 | ₦0 | 168 |
| 10 | ₦10 | ₦100 | 158 |
| 20 | ₦10 | ₦200 | 148 |
| 30 | ₦10 | ₦300 | 138 |
| 40 | ₦10 | ₦400 | 128 |
| 50 | ₦10 | ₦500 | 118 |

If we plot consumption against leisure, we would get a downward sloping line whose slope would equal the negative wage (-10). The x- intercept equals the maximum amount of leisure time Shola could enjoy: 168 hours, the number of hours in a week, the number of hours of leisure Shola would have if she didn't work at all. The y-intercept equals the maximum amount of consumption Shola could enjoy:

₦1680 a week or the amount of income Shola would earn if she worked all 168 hours of the week.

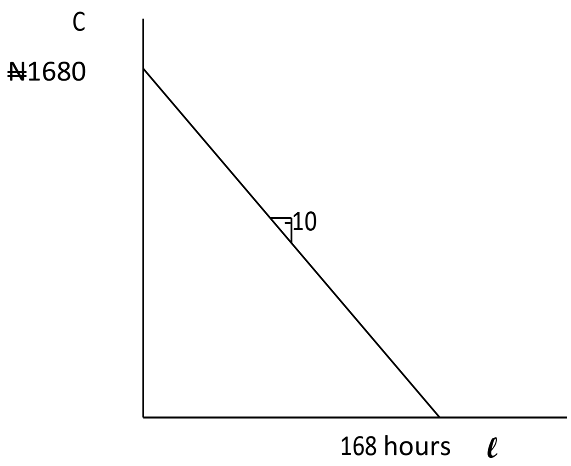



**Fig. 3.4a: Consumption against Leisure**

Every consumption-leisure combination on this line is available to Shola. Every consumption-leisure combination outside it is not available to her. Economists call this line a budget constraint. All Shola has to do to make the best labour supply decision is to pick the consumption-leisure combination on the budget constraint that she likes the best, and then work the number of hours necessary to get that combination. For example, if Shola likes ₦400 of consumption and 128 hours of leisure the best, she will want to work 40 hours a week.

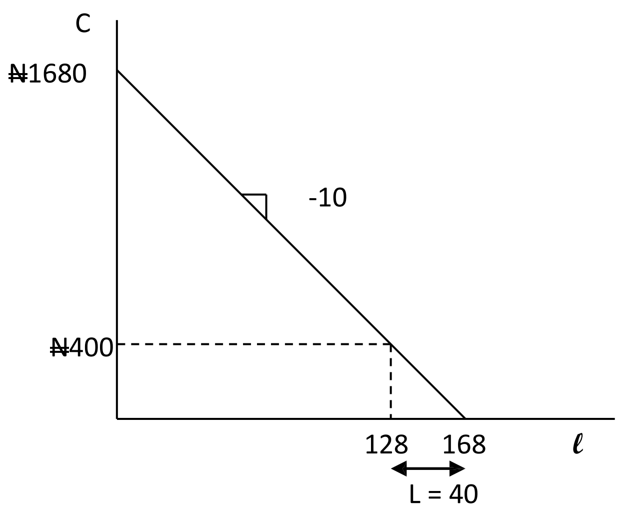


**Fig. 3.4b: Budget Constraint**

We have so far assumed that Shola earns all her income and consequently pays for all her consumption by working. But she may earn income from non-labour sources such as interest from savings accounts or bonds, dividends from stocks, welfare payments, etc. Suppose Shola earns ₦200 in non-labour income each week. This is income she earns no matter how much she works, so we can account for it by adding ₦200 in extra consumption at each level of labour supply:

**Table 3.2: Shola's Consumption-Leisure Possibilities, Non-labour Income = ₦200**

| Weekly Labour | Wage | Non-labour Income | Weekly Consumption | Weekly Leisure |
|---|---|---|---|---|
| | | | = W*L | =168 hours - L |
| 0 | ₦10 | ₦200 | ₦0 | 168 |
| 10 | ₦10 | ₦200 | ₦100 | 158 |
| 20 | ₦10 | ₦200 | ₦200 | 148 |
| 30 | ₦10 | ₦200 | ₦300 | 138 |
| 40 | ₦10 | ₦200 | ₦400 | 128 |
| 50 | ₦10 | ₦200 | ₦500 | 118 |

If we plot consumption and leisure against each other in this case, we would get a budget constraint like that at the left: a straight line coming up from the maximum amount of leisure (i.e., the amount of time available) to the level of non-labour income, then a line with a slope equal to the negative wage that intercepts the vertical axis at the maximum amount of consumption.

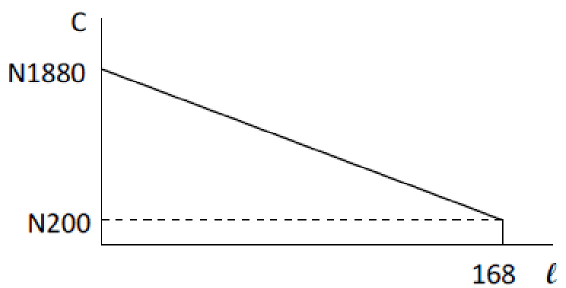



**Fig. 3.4c: Consumption-Leisure Possibilities**

We can generalise this budget constraint to the budget constraint in fig 3.4d, where w equals the employee wage, T equals the amount of time available to allocate between labour and leisure, and y equals non-labour income.

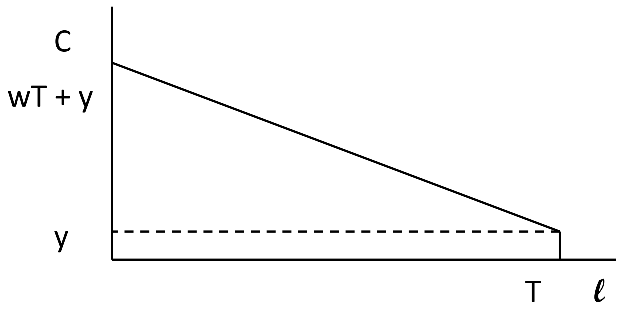



**Fig. 3.4d: Generic Budget Constraint**

The question now is; which consumption-leisure combination does Shola pick? An economist would answer that she would pick the combination that offers her the most utility. Utility is, for lack of a better synonym, an all- encompassing "goodness." It is a mythical variable that tells us just how much a consumer gets out of something. For example, if I get more enjoyment, stimulation, laughs, or what not out of an hour of watching television than out of an hour of playing football, then I get more utility out of an hour of watching television than out of an hour of playing football.

**Self-Assessment Exercise**

What is the relationship between labour supply and leisure?

### 3.3.2 Backward Bending Labour Supply

As wages increase above the subsistence level, there are two considerations affecting a worker's choice of how many hours to work per unit of time (usually day, week, or month). The first is the substitution or incentive effect. With wages rising, this says that the trade-off between working an additional hour for pay and taking one extra hour of non-paid time changes in favour of working. Thus, more hours of labour-time will be offered at the higher wage than the lower one. The second and countervailing effect is that the hours worked at the old wage rate now all gain more income than before, creating an income effect that encourages more leisure to be chosen because it is more affordable. Most economists assume that non-paid time (or "leisure") is a "normal" good which means that people want more of it as their incomes (or wealth) rise. Since a rising wage rate raises incomes, all else constant, the attraction of non-paid time rises, eventually cancelling out the substitution effect and implying the backward bending labour supply curve as indicated in fig 3.5.

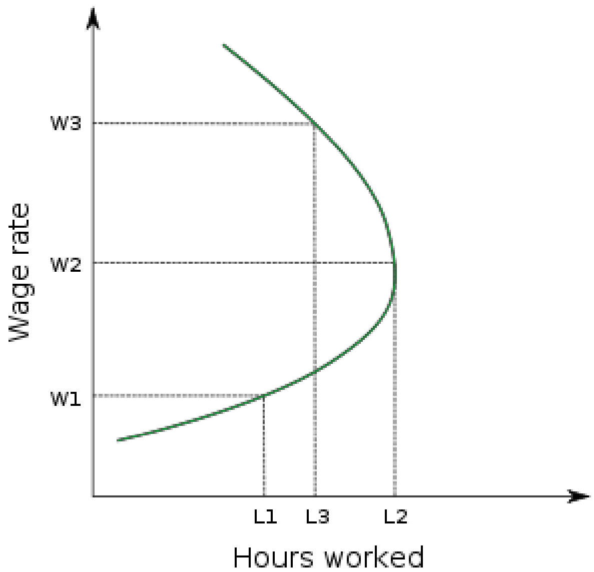



**Fig. 3.5: Backward Bend Supply**

Referring to Figure 3.5, if real wages were to increase from W1 to W2 then for an individual worker the substitution effect outweighs the income effect; therefore, they would be willing to increase their hours worked for pay from L1 to L2. However, if the real wage increased from W2 to W3, then the number of hours offered to work for pay would fall from L2 to L3. This is because the strength of the income effect now exceeds that of the substitution effect: the utility to be gained from an extra hour of non-paid time is now greater than the utility to be gained from extra income that could be earned by working the extra hour.

The above analysis only examines the effect of changing wage rates on workers already subject to those rates-that is, only these individuals' labour supply response was considered. It did not consider the additional labour supplied by workers working in other sectors (or unemployed), who are now more attracted to the jobs in the sector paying higher wages. Thus, for a given market, the wage at which the labour supply curve bends backward may be higher than the wage at which a given worker's curve bends back.

On the other hand, for the aggregate labour market, that is, a labour market without "other sectors" for workers to come from, the original story of the backward-bending labour-supply curve applies, except to the extent that some workers suffer from involuntary unemployment.

### 3.3.3 Assumptions and Caveat

**Assumptions**

1. Workers choose their hours.
2. There are no contractual obligations to work a certain number of hours.
3. Workers are utility maximising agents.
4. Work provides disutility that must be compensated for by paying wages.
5. Non-paid time is a "normal" good.

**Caveat**

Higher pay for overtime hours can reduce or negate the effect of a backward bending labour supply curve, by increasing wages only for hours worked beyond a certain amount. Overtime maintains the substitution effect at high labour supply, but the income effect from wages increasing on all previous hours worked is eliminated. Thus, higher hourly overtime pay can cause workers to work more hours than they would if the higher rate was paid on all hours.

**Self-Assessment Exercise**

Analyse the backward bending labour supply and state the assumptions and caveats

### 3.3.4 Inverted S-Shaped Labour Supply Curve

At very low wage levels-that is, near the subsistence level-the supply curve may also be curved backwards for a completely different reason. This effect creates an "inverted S" or "backward S" shape: a tail is added at the bottom of the labour-supply curve shown in figure 3.5 with the quantity of labour-time supplied falling as wages rise. In this case, because families face some minimum level of income needed to meet their subsistence requirements, lowering wages increases the amount of labour-time offered for sale. Similarly, a rise in wages can cause a decrease in the amount of labour-time offered for sale: individuals take advantage of the higher wage to spend time on needed self-or family- maintenance activities.

**Self-Assessment Exercise**

How is an inverted s-shaped labour supply curve determined?

## 3.4 Conclusion

From our discussion so far on the labour-leisure trade-off, we can infer the following:

First, consumption is increased by working more hours

Second, disadvantage of increased work hour is reduced leisure time

Third, backward bending labour supply emphasises that as wages increase above the subsistence level, substitution or income effect takes place

Fourth, higher pay for overtime hours can reduce or negate the effect of a backward bending labour supply curve.

## 4.5 Summary

Leisure is considered one good and consumption is considered another good. A consumer must make a choice between leisure and consumption. A person's consumption is the amount of labour they choose multiplied by the amount they are paid per hour of labour. The disadvantage of working more is decreased leisure and vice versa. Backward bending labour supply curve emphasise that as wages increase above the subsistence level, there are two possible outcomes which are "income and substitution effects".

## 3.6 Tutor-Marked Assignment

1. What is the effect of increased wage rate in the backward bending labour supply curve?
2. What are the assumptions and caveats of backward bending labour supply curve?
3. How is consumption-leisure combination chosen?
4. What are the causes of inverted s-shaped labour supply curve?

## 3.7 References/Further Reading

Anderson, D. & Ray, M. (2011). *Krugman's Economics for AP*. (1st ed.). New York: Worth Publishers.

Arnold, R. A. (2014). *Economics*. (11th ed.). Mason, OH: South- Western Cengage Learning.

Bade, R. & Michael, P. (2011). *Foundations of Economics, AP Edition*. (5th ed.). Upper Saddle River, NJ: Pearson.

Baumol, W. J. & Alan, S. B. (2012). *Economics: Principles and Policy*. (12th ed.). Mason, OH: South-Western Cengage Learning.

Robert, S. P. & Daniel, L. R. (2009). *Microeconomics.* (7th ed.). New Jersey: Pearson Education International.

Sloman, J. (2006). *Economics.* (6th ed,). Prentice Hall; London and New York.

### 3.8 Answers to Self-Assessment Exercises

**Question**

What is the relationship between labour supply and leisure?

**Model Answer**

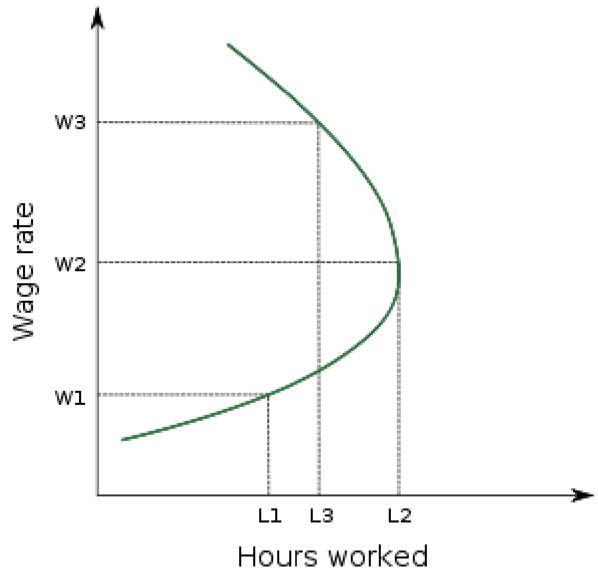
The labour supply decision is regarded as the ultimate trade-off between consumption and leisure. The advantage of working more is increased consumption; by working more hours, you earn more income and can purchase and consume more goods. The disadvantage of working more is decreased leisure; by working more hours, you have fewer hours to spend outside of work to sleep, raise children, watch TV, etc. Note that the economic definition of leisure is any time spent outside of work. Therefore, the goal of choosing how many hours you want to work-i.e., choosing your personal quantity of labour supplied-is to choose the consumption-leisure combination available to you that you like best. All the worker has to do in order to make the best labour supply decision is to pick the consumption-leisure combination on the budget constraint that he prefers best, and then work the number of hours necessary to get that combination.

**Question**

How is an inverted s-shaped labour supply curve determined?

**Model Answer**

At very low wage levels-that is, near the subsistence level-the supply curve may also be curved backwards for a completely different reason. This effect creates an "inverted S" or "backward S" shape: a tail is added at the bottom of the labour-supply curve shown on the graph below with the quantity of labour-time supplied falling as wages rise. In this case, because families face some minimum level of income needed to meet their subsistence requirements, lowering wages increases the amount of labour-time offered for sale. Similarly, a rise in wages can cause a decrease in the amount of labour-time offered for sale: individuals take advantage of the higher wage to spend time on needed self-or family- maintenance activities.


W3
W2
W1
Wage rate
L1
L3
L2
Hours worked

# Module 4 Market Structure



## Unit 1 Perfect Competition, Monopolistic Competition and Monopoly

**Contents**



### 1.1 Introduction

Market structure refers to the number of sellers and buyers in a market for a well-defined product or family of products, the distribution among the sellers or buyers of market shares, and the extent of multiplicity, differentiation, similarity or homogeneity of the products. In this module, our focus is on sellers' market structure, which refers to the number of the firms operating in the market area, the extent of product differentiation and multiplicity, and the distribution of market share among the firms. This concept of sellers' market structure is the most widely used in the characterization of market structure in the theory of the firm.

### 1.2 OBJECTIVES

At the end of this unit, you should be able to:

- explain the concept of perfect competition, monopolistic competition and monopoly

- explain both the short run and long run equilibrium in perfect and monopoly market
- advise firms in the different markets when to shut down or continue production in both the short and long run.

## 1.3 Main Content

### 1.3.1 Perfect Competition

The major characteristic of the perfect market is the existence of innumerable buyers and sellers, each independent in making decision, so that none of them can influence number of market participants or control price. The buyers or sellers in perfect competitive market are price-takers; the quantity he decides to buy or sell has no effect on price. The firm in a perfect market faces a perfectly-elastic demand curve or price line (P = AR = MR) implying that price remains constant with respect to its output (q). In other words, price will never change in response to any change in the output of the firm. Any fluctuation in the market price will be due to extraneous or exogenous factors and not because of any firm's decision. Since price (p) is constant in this respect, so also will the average revenue (AR) and marginal revenue (MR) be constant and equal. That is, p = AR = MR.

This can be demonstrated as follows:

(i) $AR = \frac{R}{q}$ but R = pq, so that $AR = \frac{pq}{q} = p$

(ii) $MR = \frac{\Delta R}{\Delta q} = \frac{R1 - R0}{q1 - q0}$

But since R1 = pq1 and R0 = pq0, MR = (pq1 – pq0) / (q1- q0)
= p(q1 – q0) / (q1 – q0) = P and MR = P = AR.

This illustrates the equilibrium condition of a profit-maximising firm in a perfectly-competitive market.

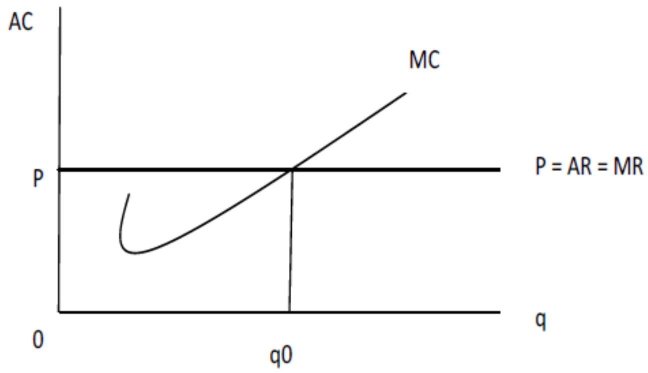



**Fig. 4.1: Profit-Maximising Equilibrium of the Firm in a Perfect Market**

## Short Run Equilibrium

The short run is the period during which some costs are fixed, particularly the fixed capital formation. Once the plant has been installed, the cost has been incurred and it cannot significantly vary as the firm varies its output within the output capacity of the plant. If the firm discovers that the investment is not worth it, it cannot throw it away as long as something can be salvaged from the investment.

In other words, if the revenue from the business covers more than the cost of variable inputs (variable cost), it does not pay to close down the business even if total cost exceeds revenue and losses are being sustained. If the firm should shut down, the loss would be fixed costs of operation. We can show this formally as:

Profit (PR) is revenue (R) – Cost (C) That is:

PR = R – C

Such that if there is a situation of loss, it is equal to C – R. Loss can be defined as

Total cost – Total Revenue

Loss = TC – TR

But, C = FC + VC, so that Loss = (FC + VC) – R

Where the firm shuts down, VC = 0, R = 0, so that Loss = FC Where the firm operates and R > VC,

Loss = FC + (VC – R) < FC Since VC – R is negative.

In other words, it pays to continue operation in the short run if the revenue more than covers the variable cost, i.e., R > VC.

For example, suppose R = 200, FC = 100 and VC = 150.

Profit = 200 – (100 + 150) = -50
Loss = 250 – 200 = 50

If the firm shuts down, R = 0, VC = 0 and Loss = FC = 100.

The loss is greater when the firm shuts down than when it continues operation because revenue is greater than variable cost. And since profit maximization implies loss minimization, the profit-maximizing firm will continue to operate in the short run whether at a loss or not, as long as revenue exceeds variable cost (R > VC).

**Long Run Equilibrium**

Since the firm intends to maximize profit, it will shut down in the long run if loss is being sustained. In the long run, all costs are variable. The firm can change its investment decisions in the long run; it can experiment with different plant sizes to search for the most suitable one; it can try various methods of production or alternative technologies. If after this entire maneuver, the firm still sustains a loss, it is predicted that it will shut down. So, the condition for continued operation in the long run is that revenue must cover all costs or profit must not be negative.

i.e., $R \geq C$ implying $AR \geq AC$

$PR = R – C \geq 0$ implies also that $R \geq C$ or $AR \geq AC$

i.e., average revenue is greater than or equal to average cost; or profit is greater than or equal to zero or not negative. This is a general condition for all profit-maximizing firms.

In particular, a firm in a perfect market cannot hold on to pure profit in the long run, since the existence of profit will attract other firms which can make their free entry into the market or industry. As many other firms enter the market, equilibrium price falls, and quantity supplied falls, average revenue falls thus reducing the profit until there is no more. For as long as there is profit, more firms will be entering the market until all profit is competed away. Long run equilibrium of a profit maximizing firm in a perfectly competitive market is given as:

MC = MR = AR = AC.

This is illustrated IN Figure 4.2.

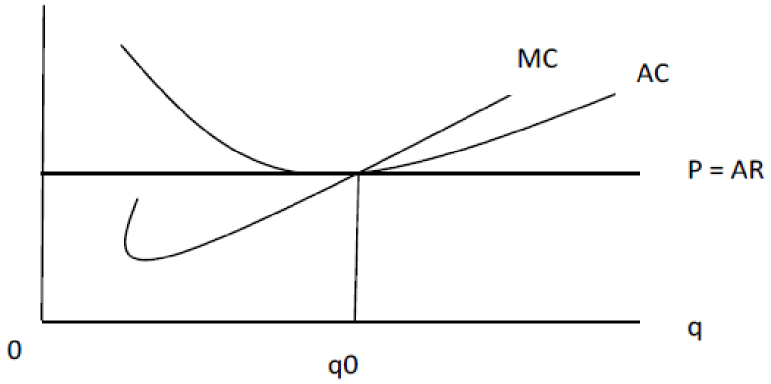



**Fig. 4.2: Long Run Equilibrium of the Profit-Maximising Firm in a Perfect Market**

In Figure 4.2, AC = AR implying zero profit: but AR = p = MR for a firm in a perfect market. While MC = MR = AR = AC. q0 is the output level that maximises profit; it is also the output that minimises average cost as well as gives the maximum and only possible output that can be produced without sustaining any loss. Thus, the goal of output maximisation and average cost minimisation coincide with that of profit maximising output.

**Self-Assessment Exercise**

Why is it difficult for a firm in a perfect market to hold unto long run pure profit?

## 1.3.2 Monopolistic Competition

Monopolistic competition is perfect competition among many firms. In other words, it is a situation where there are many firms competing with each other. However, each firm has some degree of market power; that is, each firm has some level choice over what price to charge for its products. Each of the competing firms claims to produce a different commodity whereas they produce quite similar commodities which are differentiated by brand names and other superficial means. However, because of a large number of suppliers of similar commodities, the firm in monopolistic competition faces a very elastic downward-sloping demand curve. Its features are similar to those of the monopoly or oligopoly except that there is little or no restriction to entry by new firms. This ensures that the firm in a monopolistic competition cannot hold on to profit in the long run as it will be competed away. Figure 4.3 illustrates the profit-maximising equilibrium conditions for a firm in a monopolistic competition in the long run.

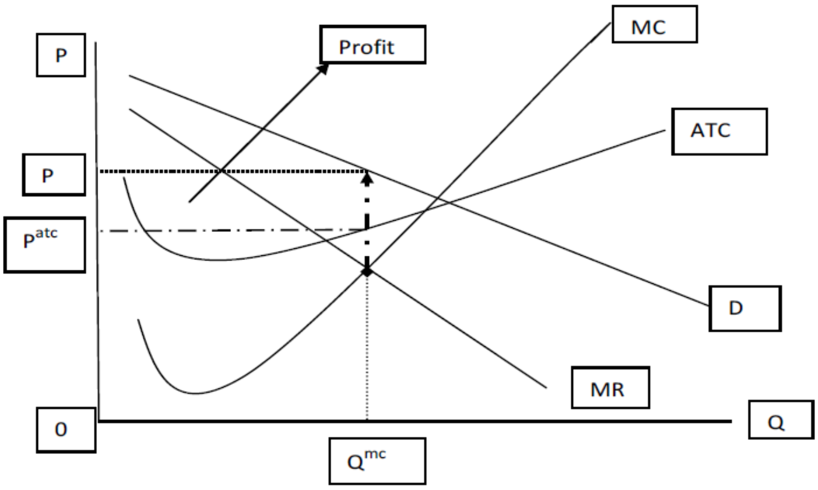



**Fig. 4.3: Profit Maximising Equilibrium of the Monopoly or Oligopoly Firm**

So, the short run and long run equilibrium condition of monopolistically competitive firm with profit maximising goal are respectively:

Short run: MC = MR; AR > AVC

Long run: MC = MR; AR = AC

It is often remarked that the firms in monopolistic competition are characterised by excess capacity as shown in Figure 4.3. At the tangency of demand curve and average cost curves, production is much below the capacity of the plant size (AC curve) or the minimum average cost (where MC cuts AC).

### 1.3.3 Monopoly

The monopoly is a situation where a single firm is the market for a given commodity or is an industry. In reality, we can speak of a monopoly firm when the supply is dominated by the firm. The firm is said to be the industry in the case of monopoly. The market is monopolized because of the existence of a barrier to entry for other firms.

**Short and Long Run Equilibrium**

The monopoly firm is capable of making losses as well as profits in the short run. When losses are sustained, the firm continues operation if AR > AVC, otherwise it shuts down, if the goal of the firm is profit maximization. The marginal revenue (MR) is also equated to marginal cost (MC) to obtain the profit-maximizing output. Where the firm makes profit, it can hold on to the profit in the long run because of barrier to competition by new firms that may desire to enter the industry.

The marginal revenue (MR) of a monopoly firm is downward sloping, unlike the perfectly-competitive firm, because the monopoly firm faces a downward-sloping demand curve.

The demand curve is the price line or the average revenue (AR) curve. Since the AR or demand is falling with output, this is because of the negative relationship between price and output, consequently the marginal revenue (MR) is below it, pulling it down. It is the general principle that when the average of any variable is falling, the marginal is below the average to pull it down, and when the average is rising, the marginal is above it to pull it up.

Figure 4.4 depicts the monopoly situation, with MR curve below AR curve, and equilibrium output (q0) determined at the interception of MC and MR curves.

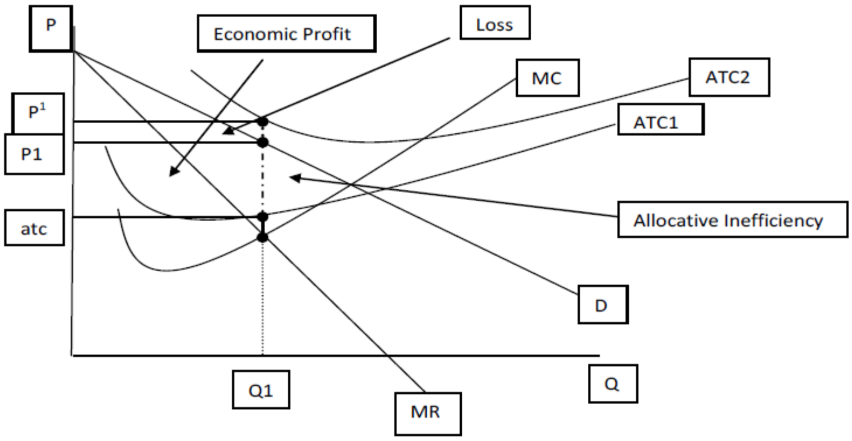



**Fig. 4.4: Profit-Maximising Equilibrium of the Firm in Monopolistic Competition**

Based on fig 4.4, the point of interception of marginal revenue (MR) and marginal cost (MC) determines the quantity supplied to the market, tracing it upward to the interception of first average total cost (ATC1) shows the cost of producing the good. Further tracing of the trend to the demand curve gives the market price. At the market price, the cost of producing the good is lower than the selling price, thus, the firm is making super-normal profit. However, the monopolist is not efficient with its production, because Allocative efficiency will occur if it produces its output at the point where marginal cost (MC) cuts the demand curve (D). The inability of the monopolist to produce at this point led to Allocative inefficiency and thus deadweight loss to the economy. Let us introduce a higher average total cost curve (ATC2), this shows that the cost of producing the output is greater than the market price, thus a loss to the monopolist. However, let's consider a situation whereby the monopolist is producing at the interception of marginal cost (MC) and demand curve (D), which is the efficient point, it was observe that the firm will still be making super-normal profit, though it is less than the initial super-normal profit.

## 1.4 Conclusion

Based on our discussion so far, the following can be inferred:

First, the firms in a perfect market face a perfectly elastic demand curve or price line P=AR=MR

Second, in the short run, if the revenue is greater than average variable cost, the firm can continue production otherwise it shuts down

Third, monopolistic competition firm is filled with many buyers and sellers selling homogenous but differentiated goods.

Fourth, the marginal revenue curve of a monopoly firm is downward sloping.

## 1.5 Summary

Market structure refers to number of buyers and sellers in a market for a well-defined product. Perfect competition is characterized by large number of buyers and sellers, and no single buyer or seller can single handedly change the market price. Equilibrium is attained when the marginal revenue is equal to marginal cost, that is, where the supply and demand curve cut each other. Monopolistic competition is similar to perfect competition, as there is no barrier to market entry. The firms sell similar but differentiated goods, and in the short-run when the average revenue is greater than the average variable cost, the firm can remain in business otherwise it shuts down. Monopoly is a situation where there is only a single seller, thus he has the power to either control the price or his output, but not the two at the same time.

## 1.6 Tutor-Marked Assignment

1. Highlight the profit maximizing condition of a monopolist.
2. What is the difference between perfect competition and monopolistic competition?
3. Discuss the minimal price making power of sellers in the monopolistic competition.

## 1.7 References/Further Reading

Chiang, E. (2014). *Core Economics*. (3rd ed.). New York: Worth Publishers.

Mansfield, E. (1975). *Microeconomics, Theory and Application.* (2nd ed.). New York: W. W. Norton and Company Inc.

Fashola, M. A. (2009). *Microeconomic Theory, Highlights & Policy Extensions for Less-Developed Economies*.

Robert, S. P. & Daniel, L. R. (2009). *Microeconomics* (7th ed.). New York: Pearson Education International.

Sloman, J. (2006). *Economics* (6th ed.). London and New York: Prentice Hall.

## 1.8 Answers to Self-Assessment Exercises

**Question**

Why is it difficult for a firm in a perfect market to hold unto long run pure profit?

**Model Answer**

A firm in a perfect market cannot hold on to pure profit in the long run, because the existence of profit will attract other firms which can make their free entry into the market or industry. As many other firms enter the market, equilibrium price falls, and quantity supplied falls, average revenue falls thus reducing the profit until there is no more. For as long as there is profit, more firms will be entering the market until all profit is competed away.

# Unit 2 Price Discrimination and Monopsony

**Contents**



## 2.1 Introduction

Price discrimination is a pricing strategy where identical or largely similar goods or services are transacted at different prices by the same provider in different markets or territories. The term differential pricing is also used to describe the practice of charging different prices to different buyers for the same quality and quantity of a product, but it can also refer to a combination of price differentiation and product differentiation. Other terms used to refer to price discrimination include equity pricing, preferential pricing and tiered pricing.

## 2.2 Objectives

At the end of this unit, you should be able to:

- define the terms price discrimination, bilateral monopoly and monopsony
- explain how price discrimination is applied in different markets
- determine the conditions for price discrimination and the type of discrimination
- analyse monopsony, price discrimination and bilateral monopoly in the Nigeria context.

## 2.3 Main Content

### 2.3.1 Price Discrimination

Price discrimination is a situation where consumers of various categories are charged different prices persistently by a firm for an identical product, or for differentiated products where the differences in production costs are negligible or insignificant relative to the differences in the prices charged. There may occur, normally and usually, some differences in the market price of the same product in different market locations or for different quantities bought by a buyer, which reflect marginal differences in production, packaging or transportation cost, and do not amount to price discrimination.

Theoretically, in a market with perfect information, perfect substitutes, and no transaction costs or prohibition on secondary exchange (or re-selling) to prevent arbitrage, price discrimination can only be a feature of monopolistic and oligopolistic markets, where market power can be exercised. Otherwise, the moment the seller tries to sell the same good at different prices, the seller at the lower price can exercise power of arbitrage by selling to the consumer buying at the higher price but with a tiny discount. However, product heterogeneity, market frictions or high fixed costs (which make marginal-cost pricing unsustainable in the long run) can allow for some degree of differential pricing to different consumers, even in fully competitive retail or industrial markets.

Price discrimination often occurs between home markets and foreign markets for a commodity produced by a firm that enjoys monopoly or remarkable market power at home. The firm may further be protected at home by import tariffs, quotas or outright ban, giving the firm undue competitive advantage in the sale of the commodity at home, while it faces perfect competition in the international market. Price discrimination usually does occur between industrial users and household consumers for various goods and services, on account of the form, specification, or standard units in which the commodity is demanded by the industrial user as opposed to the household consumer. It may also occur among various groups of industrial users or among various socioeconomic classes of household consumers, especially for social and personal services. Examples include electricity charges for industrial and household consumers, charges for utilities like gas, water and refuse disposal. Industrial products like aluminum, iron and steel, caustic soda, salt, glass, drugs and solid minerals are also sold at discriminatory prices between industrial users and household consumers, and even between small scale and large scale enterprises.

**Conditions for Price Discrimination**

It is obvious that such price discrimination will be feasible only when:

(i) the producers or sellers can control the supply to the market

(ii) the markets or buyers for which discriminatory prices are charged are or can be segregated to prevent resale between them.

For condition (i), only monopoly firm or monopoly organization of producers or sellers can control the market supply. Firms in a competitive market cannot control supply. If all producers or sellers form a cartel or a monopoly organization, they may be able to control supply.

For condition (ii), markets may be geographically separated such that resale by consumers in different markets becomes too costly due to transportation cost. Markets may also be separated geographically and politically as between two countries such that resale is impossible owing to transportation, diplomatic and political obstacles involved in crossing international boundaries as well as visa and customs regulations. Markets or buyers can also be segregated by socio-economic classes, specifically for social and personal services. Examples are first-class and second-class or economy-class categorization of the services in order to charge discriminatory prices that are not reflected in the differences in the costs of the services.

Price discrimination is profitable if feasible, given the fact that consumers have different elasticities of demand or willingness to pay. This is because profit maximization occurs at the point where marginal cost (MC) equals marginal revenue (MR). MR depends on the price elasticity of demand according to the relationship;

$$MR = P(1 - 1/\varepsilon),$$

where $\varepsilon$ is the price elasticity of demand.

If consumers with higher willingness to pay, which is reflected by the elasticity of demand, are separated from those with lower elasticity of demand, we can impose higher price on the former and impose lower price on the latter without reducing the price for the former, as opposed to a single-price monopoly where the lowest price is imposed on all, as determined by the consumer with the lowest willingness to pay.
The profit of a discriminating monopolist is the difference between his total revenue from both markets and his total cost of production. The mathematical approach to the optimizing behavior of the discriminating monopolist is presented herewith.

The profit maximizing rule is that:

Profit (Đ) = max, when MR in each market equals MC of output as a whole.
Assume a monopolist operates in two markets A and B, his cost function is given as:

$$TC = C(qA) + C(qB) = C(qA + qB) \quad (1)$$
$$TR = TRA + TRB = RA(qA) + RB(qB) \quad (2)$$

His profit function will be given as:

$$\text{Profit } (Đ) = RA(qA) + RA(qB) - C(qA + qB) \quad (3)$$

Where qA and qB are the quantities sold in the two markets and RA(qA) and RB(qB) is the revenue functions for both markets.

$$\partial Đ/\partial QA = R^{!}A(qA) - C^{!}(qA + qB) = 0 \quad (4)$$
$$\partial Đ/\partial QB = R^{!}B(qB) - C^{!}(qA + qB) = 0 \quad (5)$$
Thus $R^{!}A(qA) = R^{!}B(qB) = C^{!}(qA + qB)$
Equation (4) and (5) show that profit is maximised when

MRA = MRB = MC of output as a whole.

If the MRs are not equal, the monopolist can increase profit without affecting his TC by shifting sales from the market with low MR to the one with high MR.

## Types of Price Discrimination

### First Degree Price Discrimination

This type of price discrimination requires the monopoly seller of a good or service to know the absolute maximum price (or reservation price) that every consumer is willing to pay. In knowing the reservation price, the seller is able to sell the good or service to each consumer at the maximum price he is willing to pay, and thus transform the consumer surplus into revenues. So, the profit is equal to the sum of consumer surplus and producer surplus. The marginal consumer is the one whose reservation price equals the marginal cost of the product. The seller produces more of his product than he would to achieve monopoly profits with no price discrimination, which means that there is no deadweight loss. Examples of where this might be observed are in markets where consumers bid for tenders, though, in this case, the practice of collusive tendering could reduce the market efficiency.

### Second Degree Price Discrimination

In second degree price discrimination, price varies according to quantity demanded. Larger quantities are available at a lower unit price. This is particularly widespread in sales to industrial customers, where bulk buyers enjoy higher discounts. Additional to second degree price discrimination, sellers are not able to differentiate between different types of consumers. Thus, the suppliers will provide incentives for the consumers to differentiate themselves according to preference. As above, quantity "discounts", or non-linear pricing, is a means by which suppliers use consumer preference to distinguish classes of consumers. This allows the supplier to set different prices to the different groups and capture a larger portion of the total market surplus. In reality, different pricing may apply to differences in product quality as well as quantity. For example, airlines often offer multiple classes of seats on flights, such as first class and economy class. This is a way to differentiate consumers based on preference, and therefore allows the airline to capture greater consumers' surplus.

**Third Degree Price Discrimination**

In third degree price discrimination, price varies by attributes such as location or by customer segment, or in the most extreme case, by the individual customer's identity; where the attribute in question is used as a proxy for ability/willingness to pay. This involves charging a different price to different groups of consumers. These groups of consumers can be identified by particular characteristics such as age, sex or location.

**Self-Assessment Exercise**

What are the conditions for price discrimination?

## 2.3.2 Bilateral Monopoly

The bilateral monopoly model, with a single buyer and a single seller, can be used to analyse many types of markets, but it is most relevant for factor markets, especially those for labour services. The bilateral monopoly model was developed to explain assorted labour markets operating in the early days of the U.S. industrial revolution, in the late 1800s and early 1900s. During this period, large industrial activities (factories, mines, lumber operations) commonly created monopsony markets by dominating the labour market of a given community (a so- called company town). The expected monopsony outcome, especially low wages, inevitably resulted.

The workers sought to counter these less than desirable situations, by forming labour unions. The expressed goal of most unions was to monopolize the selling side of a labour market and balance the monopsony power of the employer. This resulted in a bilateral monopoly. Bilateral monopoly is a market containing a single buyer and a single seller, or the combination of a monopoly market and a monopsony market. A market dominated by a profit-maximizing monopoly tends to charge a higher price. A market dominated by a profit-maximizing monopsony tends to pay a lower price.

When combined into a bilateral monopoly, the buyer and seller both cannot maximize profit simultaneously and are forced to negotiate a price and quantity. The resulting price could be anywhere between the higher monopoly price and the lower monopsony price. The final price depends on the relative negotiating power of each side.

**Modern Example of Monopsony**

Markets dominated by a company-town-type monopsony employer on the buying side and a monopoly labour union on the selling side are rare in the modern economy, but the bilateral monopoly model provides insight into other markets. Consider the following:

**Industrial Labour:** While the company town might be largely a thing of the past, a number of modern labour markets contain a dominant employer on the buying side and a dominant labour union on the selling side. Negotiations between large corporations and labour unions over wages and working conditions approximate the bilateral monopoly model. While these markets are not, strictly speaking, comprised of pure monopoly sellers and pure monopsony buyers, they do have a definite bilateral monopoly flavor.

**Consumer Goods**: A number of modern consumption purchases, such as buying a house or a car, tend to fit the bilateral monopoly model. These purchases often include a single buyer and a single seller. As the bilateral model suggests, one-on-one negotiation over the price is a common feature of such purchases.

**Athletic Labour Services**: Perhaps the factor markets that come closest to the bilateral monopoly model are those involving the labour services of professional athletes. Players in each league are also represented by a players' union, including the National Football League Players Association. While the wage rates tend to be significantly higher for athletes than for factory workers, the underlying process is much the same.

**Self-Assessment Exercise**

What do you understand by the term bilateral monopoly?

### 2.3.3 Monopsony

Monopsony is a market in which a single buyer completely controls the demand for a good. While the market for any type of good, service, resource, or commodity could, in principle, function as monopsony, this form of market structure tends to be most pronounced for the exchange of factor services. While the real world does not contain monopsony in its absolute purest form, labour markets in which a single large factory is the dominate employer in a small community comes as close as any. Like a monopoly seller, a monopsony buyer is a price maker with complete market control. Monopsony is also comparable to monopoly in terms of inefficiency. Monopsony does not generate an efficient allocation of resources. The price paid by a monopsony is lower and the quantity exchanged is less than with the benchmark of perfect competition.

**Characteristics**

The three key characteristics of monopsony are:

(1) a single firm buying all output in a market

(2) no alternative buyers

(3) restrictions on entry into the industry.

**Single Buyer:** First and foremost, a monopsony is a monopsony because it is the only buyer in the market. The word monopsony actually translates as "one buyer." As the only buyer, a monopsony controls the demand-side of the market completely. If anyone wants to sell the good, they must sell to the monopsony.

**No Alternatives:** A monopsony achieves single-buyer status because sellers have no alternative buyers for their goods. This is the key characteristic that usually prevents monopsony from existing in the real world in its pure, ideal form. Sellers almost always have alternatives.

**Barriers to Entry:** A monopsony often acquires, and generally maintains single buyer status due to restrictions on the entry of other buyers into the market. The key barriers to entry are much the same as those that exist for monopoly:

(1) government license or franchise

(2) resource ownership

(3) patents and copyrights

(4) high start-up cost

(5) decreasing average total cost.

**Self-Assessment Exercise**

What are the characteristics of monopsony?

## 2.4 Conclusion

We conclude with the following issues from our discussion above:

First, a monopsony buyer is a price maker, and does not generate an efficient allocation of resources.

Second, for price discrimination to occur, the producer or seller must control the supply to the market, and be sure of segregated markets in other to avoid resale between them.

Third, bilateral monopoly is a market for a single buyer and seller or the combination of a monopoly market and a monopsony market.

Fourth, profit cannot be made by both the buyers and sellers simultaneously in the bilateral monopoly market, so they are both forced to resort to negotiation on price and quantity, with the winner being the better negotiator.

## 2.5 Summary

Price discrimination can only be a feature of monopolistic and oligopolistic markets where there is perfect information, perfect substitutes and no transaction costs or prohibition on secondary exchange to prevent arbitrage. Price discrimination often occurs between the home markets and foreign markets for a commodity produced by a firm that enjoys monopoly or remarkable market power at home. Bilateral monopoly is a market containing a single buyer and a single seller, or the combination of a monopoly market and a monopsony market, as a market dominated by a profit maximizing monopoly tends to charge a higher price, and a market dominated by a profit maximizing monopsony tends to pay a lower price, thus when combined into a bilateral monopoly, the buyer and the seller both cannot maximize profit simultaneously and are forced to negotiate a price and quantity. Monopsony is where a single buyer completely controls the demand for a good. A monopsony buyer is a price maker with complete market control.

## 2.6 Tutor-Marked Assignment

1. Differentiate between the different types of price discrimination.
2. Analyse how the single buyers and sellers interact in a bilateral monopoly market.
3. Give an example of a monopsony market in Nigeria.

## 2.7 References/Further Reading

Case, et al. (2014). *Principles of Economics*. (11th ed.). Upper Saddle River, New Jersey: Pearson.

Chiang, E. (2014). *Core Economics*. (3rd. ed.). New York: Worth Publishers.

Colander, D. C. (2013). *Economics*. (9th ed.). New York: McGraw-Hill.

Fashola, M. A. (2009). *Microeconomic theory, Highlights & Policy Extentions for Less-Developed Economies.*

Robert, S. P. & Rubinfeld, D. L. (2009). *Microeconomics.* (7th ed.). New York: Pearson Education International.

## 2.8 Answers to Self-Assessment Exercises

**Question**

What do you understand by the term bilateral monopoly?

**Model Answer**

Bilateral monopoly is a market containing a single buyer and a single seller, or the combination of a monopoly market and a monopsony market. A market dominated by a profit-maximising monopoly tends to charge a higher price. A market dominated by a profit-maximising monopsony tends to pay a lower price. When combined into a bilateral monopoly, the buyer and seller both cannot maximise profit simultaneously and are forced to negotiate a price and quantity. The resulting price could be anywhere between the higher monopoly price and the lower monopsony price. The final price depends on the relative negotiating power of each side.

# Unit 3 Collusive Oligopoly and Non- Collusive Oligopoly

**Contents**



## 3.1 Introduction

Oligopoly consists of a few dominant firms in the industry. This means that the number of firms in the industry may be small or large provided that a few (four to nine) dominate the market or control more than 50% of market share. Concentration ratio is often used to determine whether the marker is an oligopoly. When the number of firms is very small, the oligopoly market may be capable of engaging in collusion deliberately or tacitly. When the number is not quite few, the firms will tend to behave as if they are in a perfectly competitive market by assuming that prices will remain constant as each firm alters its output.

## 3.2 Objectives

At the end of this unit, you should be able to:

- analyse the Bertrand, Cournot, Stackelberg, Edgeworth and the Chamberlin oligopolistic model
- state the reactions and functions of rival firms in an oligopolistic market
- determine the profit condition of oligopolistic firms
- explain the effect of price and output variation on profit.

## 3.3 Main Content

### 3.3.1 Collusive and Non-Collusive Oligopoly

**The Bertrand Model**

This model addresses the case of not so many number of firms in the market, such that each firm believes or assumes that its activity will not alter the market price for the product. So, the firm then attempts to maximise its profit, holding the price constant. In this case, the marginal revenue is constant and equal to the price. The profit maximising equilibrium of the firm is then given by MC = P. This oligopoly model tends to produce minimum profit and the largest output among all types of oligopoly models, as illustrated in Figure 4.5

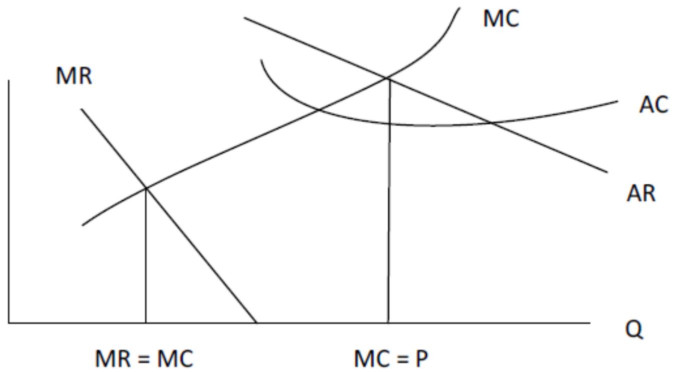



**Fig. 4.5: Equilibrium of Bertrand oligopoly model at MC=P**

The monopoly output is much lower at MC = MR

**Self-Assessment Exercise**

Explain the Bertrand's oligopoly model

**The Cournot Classical Model**

In the Cournot classical model, the firm assumes that other firms will not change their output as it changes its own output. On that count, the firm maximises profit by equating the marginal cost (MC) to its marginal revenue (MR). However, the reality is that other firms will change their output and this forces price and the marginal revenue of the firm down, requiring the firm to reduce its output as a reaction to the other firms' action. So, initially, the firms will tend to produce more than the collusion oligopoly and realise less profit. They will however achieve equilibrium eventually. This can be illustrated by the intersection of reaction curves of the firms in a duopoly model.

P = f(Q1, Q2)

Profits for the firms are respectively:

P1 = R1 – C1, where R1 = f(Q1), and C1 = f(Q1)
P2 = R2 – C2, where R2 = f(Q2), and C2 = f(Q2)

Since revenue
R1 = PQ1 = f(Q1, Q2)Q1 and R2 = PQ2 = f(Q1, Q2)Q2,

The reaction functions obtained by taking the partial derivatives of profits of the firms and equating it to zero, produces two equations in two unknown, Q1 and Q2, with definite solutions. The output so derived produce small outputs than with Bertrand oligopoly model, but larger outputs than what is produced with the collusion model. While the profits are larger than with the Bertrand model, but smaller than what is obtainable with the collusion model.
This can be illustrated with a numerical example.

Suppose price (P) = 300 – 2(Q1 + Q2),

Cost (C1) = 8Q1, and C2 = 10Q2, profit functions are respectively:

$$Đ1 = R1 - C1 = PQ1 - C1 = [300 - 2(Q1 + Q2)]\ Q1 - 8Q1$$
$$= 300Q1 - 2Q1^2 - 2Q1Q2 - 8Q1$$

$$Đ2 = R2 - C2 = PQ2 - C2 = [300 - 2(Q1 + Q2)]\ Q2 - 10Q2$$
$$= 300Q2 - 2Q2^2 - 2Q1Q2 - 10Q2$$

The reaction functions are then as follows:

$$\partial Đ1/\partial Q1 = 300 - 4Q1 - 2Q2 - 8 = 0$$

$\partial Đ2/\partial Q2 = 300 - 4Q2 - 2Q1 - 10 = 0$

The solution is given by Q1 = 49 and Q2 = 48. This could be checked and it will be observed that it produces smaller profit and larger total output if the total profit of the oligopoly firms were to be maximized in a collusion model.

**Assumptions of Cournot Model**

1. Cournot assumes the existence of a costless monopolist whose market is penetrated by a second firm with identical cost conditions. This results in the existence of a duopoly situation consisting of firms I and II.
2. Each firm produces identical products.
3. There is perfect flow of information. Each firm is fully aware of the demand curve facing the market and identical prices.
4. It is a costless monopoly, as MC = 0
5. Each duopolist acts independently (absence of collusion) under the assumption that rivals output will remain exactly where it is now even if firm I changes its output. While this assumption is rather old and naïve, Figure 4.6 demonstrates the simplified Cournot model.

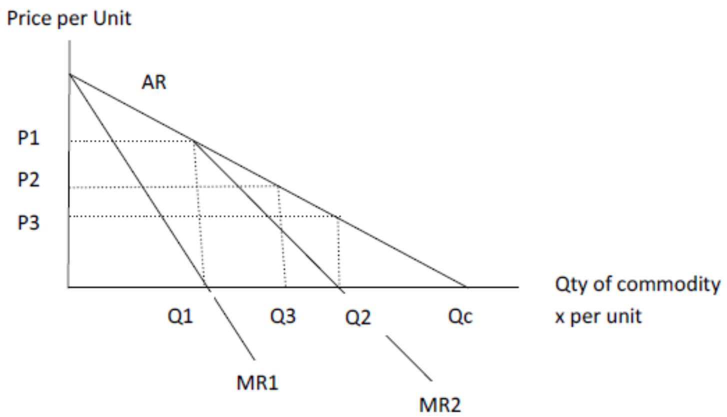



**Fig. 4.6: The Cournot Model**

Assume Firm I entered the market first with market demand and marginal revenue curves represented as AR and MR1 respectively. Operating as a pure monopolist, he will produce an equilibrium output 0Q1 where MR=MC=0, and sell at a per unit price of OP1. This is a monopoly price and the firm secures a monopoly profit. Note that if pure competition prevailed, the market equilibrium output would be 0Qc since MC=0 and P=0.

Now, assume that firm II enters the market believing that firm I's output will be invariant at 0Q1, he regards AQc as the remaining market demand. Introducing his MR curve (MR2) from point A, the profit maximising output for firm II (Q1Q) is half the remaining market (AQc) which is one fourth of the entire market. He then equates his MR and MC at output level 0Q2. His price is 0P2. Both firms now sell at the price 0P2.

Firm I in turn assumes that firm II's output will remain fixed at Q2C2 which is ¼ of the entire market. Thus, the total market demand he now faces becomes 3/4., which is equivalent to 0Q2. He now proceeds to maximise his profit by producing half of the market which is ½ of ¾ of the market. This process continues until a limit or convergence point is reached. In equilibrium, the two duopolists will produce two third of the competitive output (2/3 Qc) and charge two third of the monopoly price (2/3 P1). Their joint profit is 2/3 of the monopoly profit.

**Self-Assessment Exercise**

What are the central assumptions o f the Cournot approach to oligopolistic market conduct?

**Edgeworth's Model**

An English economist Francis Edgeworth developed a model similar to Bertrand's in the assumption of zero conjectural variation in prices. He however differs from Bertrand in assuming that the two firms have limited capacity to produce OQc units at P=0, (Figure 4.6), then a lower limit could be established beyond which price could not fall. But such a price would be unstable. However, once it is reached; one of the rivals will assume the other will stick to that price where it reaches its capacity limit. It then perceives that at such a point:

(i) The rival firm cannot sell more because it has reached its capacity limit.

(ii) An increase in price will increase profits.

Consequently, one of the firms will raise its price to a point where it maximises its profits in the remaining markets. The other will presumably follow suit, but to a level slightly below that of the first firm. The first firm naturally reacts with another price cut. As the process continues, the price level falls back to the lower limit. At this point, the opportunity for any of the firms to gain once more by raising price recurs and price goes up again. This Edgeworth proposal, goes on and on, as long as each firm adheres to the assumption of zero conjectural variation in prices on the part of rival firms.

**Self-Assessment Exercise**

What are the major differences between the Cournot model and the Edgeworth model?

**Stackelberg Model**

This model gives comprehensive possibilities to oligopoly market behaviour by assuming realistically that the firm in an oligopoly market is aware that its profits and prospects depend not only on its own output decisions but also on the decisions of rival firms about the level of their output, among others. Thus, the model recognises oligopolistic interdependence and its influence on output and price decisions of the firm. This specifically means that the behaviour of the oligopoly firm in respect of output or price decisions takes into account the reactions of rival firms. Although, Stackelberg assumes two firms in the market, the analysis may be extended to more than two firms.

Suppose "n" represents the number of firms in the oligopoly market, and then the profit of each firm "I" depends directly on its own output levels and that of other firms. Thus, given that:

$Qi = fi(Q1, Q2, Q3, \ldots Qn); I = 1, 2, 3, \ldots, n,$

the equilibrium of the profit maximising oligopoly firms is respectively given by:

$$\frac{\partial \pi 1}{\partial Q1} = \frac{\partial f1}{\partial Q1} + \frac{\partial f1}{\partial Q2} * \frac{\partial Q2}{\partial Q1} + \frac{\partial f1}{\partial Q3} * \frac{\partial Q3}{\partial Q1} + \ldots + \frac{\partial f1}{\partial Qn} * \frac{\partial Qn}{\partial Q1} = 0 \qquad (1)$$

$$\frac{\partial \pi 2}{\partial Q2} = \frac{\partial f2}{\partial Q2} + \frac{\partial f2}{\partial Q1} * \frac{\partial Q1}{\partial Q2} + \frac{\partial f2}{\partial Q3} * \frac{\partial Q3}{\partial Q2} + \ldots + \frac{\partial f2}{\partial Qn} * \frac{\partial Qn}{\partial Q2} = 0 \qquad (2)$$

$$\frac{\partial \pi n}{\partial Qn} = \frac{\partial fn}{\partial Qn} + \frac{\partial fn}{\partial Q1} * \frac{\partial Q1}{\partial Qn} + \frac{\partial fn}{\partial Q2} * \frac{\partial Q2}{\partial Qn} + \ldots + \frac{\partial fn}{\partial Qn-1} * \frac{\partial Qn-1}{\partial Qn} = 0 \qquad (3)$$

Restricting the analysis as presented by Stackelberg to two oligopoly firms (A and B), the profit function is given by:

$$\frac{\partial \pi A}{\partial QA} = \frac{\partial fA}{\partial QA} + \frac{\partial fA}{\partial QB} * \frac{\partial QB}{\partial QA} = 0 \qquad (4)$$

$$\frac{\partial \pi B}{\partial QB} = \frac{\partial fB}{\partial QB} + \frac{\partial fB}{\partial QA} * \frac{\partial QA}{\partial QB} = 0$$

(5)

The cross partial derivatives ($\partial Q_B/\partial Q_A$) and ($\partial Q_A/\partial Q_B$) relating the output of firm B to that of firm A and vice versa, respectively, in terms involving the chain rule in the system (4) describe the reaction of firm B to output decision of firm A, and the reaction of firm A to output decision of firm B, while $\frac{\partial fA}{\partial QB}$ and $\frac{\partial fB}{\partial QA}$ describe the profit maximising response of firm A to the output expansion of firm B, and vice versa. The problem with this model is that each firm may not be able to predict the reaction to its output decisions by the rival firms. And that is why the model of Bertrand and Cournot made the simple assumption that the oligopoly firm would, in the pursuit of its profit-maximizing goal, disregard the reaction of the rival firms in its possible impact on the market prices or aggregate.

To address the problem of the inability of firms to predict accurately, Stackelberg proposed four possibilities.

1. Firm A plays leadership role by being in a position to know or dictate the reaction of firm B that plays the followership role. Firm B, on the other hand, would react to the decision of firm A according to the understanding

between them. In that case, the simultaneous system has a solution known as Stackelberg equilibrium. Thus, the cross partial derivative ($\partial Q_B/\partial Q_A$) vanishes while the cross partial derivative ($\partial Q_A/\partial Q_B$) has a known functional value.

2. The above solution is also true in a case where firm B plays leadership role while firm A plays the followership role.
3. This is when both firm play followership roles, each expecting that the other firm will not at all react to the decision taken by them. Thus, the term involving the cross partial derivatives ($\partial Q_B/\partial Q_A$) and ($\partial Q_A/\partial Q_B$) vanishes, and so do the terms involving the chain rule. So, the system reduces to that of the Cournot model.
4. Both firms are unable to predict the reaction of the other firms, and so the term involving the cross partial derivatives in the chain rule have an unknown functional value. There will therefore be no solution to the system, and we then have Stackelberg disequilibrium. The two firms are unable to reach an understanding in their behaviour and this is tantamount to each firm asserting its independence, or both playing leadership roles as Stackelberg puts it.

**Self-Assessment Exercise**

How can you predict the reaction of a rival firm under Stackelberg model?

**The Collusion Model**

The Chamberlin's small group oligopoly model consists of very few firms, and permission of collusion without open agreement but with tacit understanding. This leads to a monopoly solution with maximum total profit and minimum total output. The equilibrium in such a model is that maximization of total profit requires that marginal total profit with respect to the output of every firm be equated to zero (i.e. $\partial P/\partial Qi = 0$; I = 1, 2, . . ., n).

Tacit collusion can be obtained through many strategies: One is the price leadership, where the leading firm fixes price and others follow, and so competition is largely non-price but by advertisements and other methods. Another strategy is sharing of markets by regions, by quotas in an informal manner. Tacitly, also, firms may not engage in price war and each operates to stabilize price and only compete through other strategies like advertisement. The kinked demand model is consistent with the collusion model as with other models.

**Self-Assessment Exercise**

Given that:

Profit = 300 – 0.5Q

Where Q = QA + QB

$CA = 4000 – 20QA + 0.2QA^2$ and

$CB = 3000 + 15B + 0.25QB^2$

Calculate the total profit.

**The Kinked Demand Model**

This model postulates that oligopoly firms realize their interdependence such that they will avoid price competition below a particular level. If any firm cuts price below that level (or kinked), others will react by cutting their own price, so that the first to cut price will not divert customers in a significant manner to the firm whose price has been cut. With the reaction of others, the firm will not gain any significant number of customers. Therefore, the demand curve below shows that price level (the kink) is much more inelastic than the segment of the demand curve above the kink. This illustration is presented in Figure 4.7.

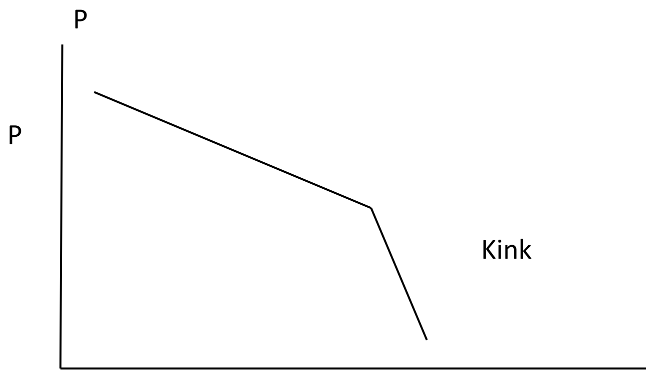



**Fig. 4.**7: **The Kinked Demand Curve**

If any firm attempts to raise its price, others are not likely to follow, with the consequence that the firm will lose a lot of customers to its competitors. So, there is a tendency for prices to cluster at or a little above the kink. Except when there is a case of price leadership by a dominant firm, price increase by one firm will not be followed by others. When there is a recognized leader and price leadership exists, a price increase by the dominant firm may be followed by the others, leading to a new kink on the demand curve.

**Self-Assessment Exercise**

Is it possible for an oligopolistic firm confronted with a kinked demand curve situation to increase its short run profits by increasing its short run expenditure on sales promotion?

## 3.4 Conclusion

We have so far examined the classical solution of Cournot, Chamberlin. *et al* to the oligopolistic market behaviour. These theories explained the reason for price rigidity in an oligopolistic industry, and why prices change simultaneously throughout the industry whenever it occurs. We then discussed how oligopolistic firms decide upon a particular price.

## 3.5 Summary

Based on the discussion on Collusive and Non-Collusive Oligopoly, the following can be inferred:

First, Kinked demand is a model of steady prices in non-collusive oligopolistic market

Second, an oligopolist firm will refrain from independent initiating price cut because this will elicit prompt reaction from rival firms with similar cuts, thus wiping off any anticipated gains in profits and market share position.

Third, the price charged by an oligopolist firm tends to be rigid but the output however tends to respond to moderate changes in market demand conditions.

Fourth, Oligopolist firms cannot coexist by charging substantially different prices for essentially similar commodities.

Fifth firms with higher prices will price themselves out of business as lower priced competitors make the sale. Consequently, prices tend to remain steady in oligopolistic markets.

Finally, the price and quantity produced per unit of time by an oligopolistic firm tend to remain rigid in the face of moderate cost changes.

## 3.6 Tutor-Marked Assignment

1. Consider an oligopoly market with two firms A and B, facing the demand function:

   $P=300 – 0.5Q$

   Suppose their cost functions are respectively:

   $C_A = 4000 + 20Q_A + 0.2Q_A$ and $C_B = 3000 + 15Q_B + 0.25Q_B^2$

   Derive the reaction functions of the firm and the profit maximising equilibrium under Cournot (classical oligopoly), Bertrand and the Collusive model.

2. What are the essential differences between the Cournot, Edgeworth, Stackelberg, Chamberlin and Bertrand Model?

## 7.0 References/Further Reading

Robert, S. P. & Daniel, L. R. (2009). *Microeconomics.* (7th ed.). Pearson Education International.

Koutsoyiannis, A. (1987). *Modern Microeconomics.* (2nd ed.).

Mansfield, Edwin. (1975). *Microeconomics, Theory and Application* (2nd ed.). New York: W. W. Norton and Company Inc.

Fashola, M. A. (2009). *Microeconomic theory: Highlights & Policy Extensions for Less-Developed Economies.*

Lipsey, R. G. & Chrystal, K.A. (2006). *Economics.* (10th ed.). New Delhi, Oxford University Press.

Samuelson, P.A. & Nordhaus, W.D. (2002). *Economics.* ($17^{th}$ ed,). New Delhi, Tata McGraw-Hill edition.

## 3.8 Answers to Self-Assessment Exercises

**Question**

What are the central assumptions o f the Cournot approach to oligopolistic market conduct?

**Model Answer**

The central assumptions of the Cournot approach to oligopolistic market are:

1. Cournot assumes the existence of a costless monopolist whose market is penetrated by a second firm with identical cost conditions. This results in the existence of a duopoly situation consisting of firms I and II.
2. Each firm produces identical products.
3. There is perfect flow of information. Each firm is fully aware of the demand curve facing the market and identical prices.
4. It is a costless monopoly, as MC = 0
5. Each duopolist acts independently (absence of collusion) under the assumption that rivals output will remain exactly where it is now even if firm I changes its output.

**Question**

How can you predict the reaction of a rival firm under Stackelberg model?

**Model Answer**

Under the Stackelberg model, there are four possible ways by which a firm can predict the reaction of a rival firm.

1. Firm A plays leadership role by being in a position to know or dictate the reaction of firm B that plays the followership role. Firm B, on the other hand, would react to the decision of firm A according to the understanding between them. In that case, the simultaneous system has a solution known as

Stackelberg equilibrium. Thus, the cross partial derivative ($\partial Q_B/\partial Q_A$) vanishes while the cross partial derivative ($\partial Q_A/\partial Q_B$) has a known functional value.

2. The above solution is also true in a case where firm B plays leadership role while firm A plays the followership role.
3. This is when both firm play followership roles, each expecting that the other firm will not at all react to the decision taken by them. Thus, the term involving the cross partial derivatives ($\partial Q_B/\partial Q_A$) and ($\partial Q_A/\partial Q_B$) vanishes, and so do the terms involving the chain rule. So, the system reduces to that of the Cournot model.
4. Both firms are unable to predict the reaction of the other firms, and so the term involving the cross partial derivatives in the chain rule have an unknown functional value. There will therefore be no solution to the system, and we then have Stackelberg disequilibrium. The two firms are unable to reach an understanding in their behaviour and this is tantamount to each firm asserting its independence, or both playing leadership roles as Stackelberg puts it.

**Question**

Is it possible for an oligopolistic firm confronted with a kinked demand curve situation to increase its short run profits by increasing its short run expenditure on sales promotion?

**Model Answer**

It is possible for an oligopolistic firm confronted with a kinked demand curve situation to increase its short run profits by increasing its short run expenditure on sales promotion. This is essentially because firms in oligopolistic market confronted with kinked demand curve tacitly do not engage in price war or competition. They rather, operate in a manner that stabilizes price and only compete through other strategies such as advertisement. Thus, the kinked demand model is consistent with the collusion model.

# Module 5 Linear Programming, General Equilibrium and Social Welfare



## Unit 1 Linear Programming

**Contents**



## 1.1 Introduction

The linear programming problem is that of choosing nonnegative values of certain variables so as to maximise or minimise a given linear function subject to a given set of linear inequality constraints. It can also be referred to as the use of linear mathematical relations to plan production activities. Linear programming is a resource allocation tool in production economics.

Linear programs are problems that can be expressed in canonical form:

$$\begin{aligned} &maximise\ c^T x \\ &subject\ to\ Ax \le b \\ &maximise\ x \ge 0 \end{aligned}$$

where **x** represents the vector of variables (to be determined), **c** and **b** are vectors of (known) coefficients, *A* is a (known) matrix of coefficients, and $(\cdot)^T$ is the matrix transpose. The expression to be maximized or minimized is called the *objective function* ($\mathbf{c}^T\mathbf{x}$ in this case). The inequalities $A\mathbf{x} \leq \mathbf{b}$ and $\mathbf{x} \geq \mathbf{0}$ are the constraints which specify a convex polytopeover which is the objective function to be optimised. In this context, two vectors are comparable when they have the same dimensions. If every entry in the first is less- than or equal-to the corresponding entry in the second, then we can say the first vector is less-than or equal-to the second vector.

Linear programming can be applied to various fields of study. It is used in business and economics, but can also be utilised for some engineering problems. Industries that use linear programming models include transportation, energy, telecommunications, and manufacturing. It has proved useful in modeling diverse types of problems in planning, routing, scheduling, assignment, and design.

Linear programming maximum problem looks like the equation presented below:

Max $f(x) = cx$

Subject to $Ax \leq b$, $x \geq 0$.

More generally,

Max $Z = P_1X_1 + P_2X_2$

Subject to

$a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n \leq b1$

$a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n \leq b2$

$a_{m1}x_1 + a_{m2}x_2 + \dots + a_{mn}x_n \leq bm$

$x_1 \geq 0, x_2 \geq 0, \dots, x_n \geq 0.$

Where $a_{11}$is the amount of resources required to produce a unit of activity, i.e., the recommended unit of output to produce specified quantity of output (Linear programming has to do with using the optimal level of input to achieve an optimum number of output).

## 1.2 Objectives

At the end of this unit, you should be able to:

- state the assumptions of linear programming
- state the advantages and disadvantages of linear programming
- explain the standard form canonical tableaux
- solve maximisation and minimisation problem.

## 1.3 Main Content

### 1.3.1 Assumptions of LP for Production

1. The objective function is linear
2. There is an additivity of resources and activities
3. Activities and resources are non-negative
4. Resources and activities are infinitely divisible
5. Resources and activities are finite
6. There is a linear relationship between activities and resources
7. Linear programming models are deterministic and not stochastic.

### 1.3.2 Advantages and Limitations

**Advantages**

i. It helps decision-makers to use their productive resources effectively.

ii. The decision-making approach of the user becomes more objective and less subjective.

iii. In a production process, bottlenecks may occur. For example, in a factory, some machines may be in great demand, while others may lie idle for some times.

iv. A significant advantage of linear programming is highlighting such bottle necks.

**Limitations**

i, Linear programming is applicable only to problems where the constraints and objective function are linear i.e., where they can be expressed as equations which represents straight lines. In real life situations, when constraints or objective functions are not linear, this technique cannot be used.

ii. Factors such as uncertainty and time are not taken into consideration.

iii. Parameters in the model are assumed to be constant but in real life situations, they are not constant.

iv. Linear programming deals with only single objectives, whereas in real life situations, we may have multiple and conflicting objectives.

v. In solving linear programming problem, there is no guarantee that we will get an integer value. In some cases of no men/machine, a non-integer value is meaningless.

**Self-Assessment Exercise**

What are the advantages and disadvantages of linear programming?

### 1.3.3 Construction of the Model

The transformation of a linear programme to one in standard form may be accomplished as follows. First, for each variable with a lower bound other than 0, a new variable is introduced representing the difference between the variable and the bound. The original variable can then be eliminated by substitution. For example, given the constraint:

$X_1 \geq 5$

a new variable, y1 is introduced with $y_1 = x_1 - 5$

$x_1 = y_1 + 5$

The second equation may be used to eliminate x1 from the linear programme. In this way, all lower bound constraints may be changed to non-negative restrictions. Second, for each remaining inequality constraint, a new variable, called a slack variable, is introduced to change the constraint to an equality constraint. This variable represents the difference between the two sides of the inequality and is assumed to be nonnegative. For example, the inequalities:

$x_2 + 2x_3 \leq 3$
$-x_4 + 3x_5 \geq 2$

Are replaced with

$x_2 + 2x_3 + s_1 = 3$
$-x_4 + 3x_5 + s_2 = 2$
$s_1, s_2 \geq 0$

It is much easier to perform algebraic manipulation on inequalities in this form. Third, each unrestricted variable is eliminated from the linear program. This can be done in two ways; one is by solving for the variable in one of the equations in which it appears and then eliminating the variable by substitution. The other is to replace the variable with the difference of two restricted variables. For example, if z1 is unrestricted then write:

$z_1 = z_1^+ - z_1^-$

$z_1^+, z_1^- \geq 0$

The equation may be used to eliminate z1 from the linear program. When this process is complete the feasible region will be in the form:
$Ax = b, xi \geq 0$

### 1.3.4 Canonical Tableaux

A linear programme in standard form can be represented as a tableau of the form:

$$\begin{bmatrix} 1 & -\mathbf{c}^T & 0 \\ 0 & \mathbf{A} & \mathbf{b} \end{bmatrix}$$

The first row defines the objective function and the remaining rows specify the constraints. If the columns of A can be rearranged so that it contains the identity matrix of order p (the number of rows in A) then the tableau is said to be in canonical form. The variables corresponding to the columns of the identity matrix are called basic variables while the remaining variables are called non-basic or free variables. If the non- basic variables are assumed to be 0, then the values of the basic variables are easily obtained as entries in b and this solution is a basic feasible solution.

Conversely, given a basic feasible solution, the columns corresponding to the non-zero variables can be expanded to a non-singular matrix. If the corresponding tableau is multiplied by the inverse of this matrix, then the result is a tableau in canonical form.

$$\begin{bmatrix} 1 & -\mathbf{c}_B^T & -\mathbf{c}_D^T & 0 \\ 0 & I & \mathbf{D} & \mathbf{b} \end{bmatrix}$$

Let the expression below be a tableau in canonical form. Additional row-addition transformations can be applied to remove the coefficients $c^T{}_B$ from the objective function. This process is called pricing out and results in a canonical tableau where zB is the value of the objective function at the corresponding basic feasible solution. The updated coefficients, also known as relative cost coefficients, are the rates of change of the objective function with respect to the non-basic variables.

$$\begin{bmatrix} 1 & 0 & -\bar{\mathbf{c}}_D^T & z_B \\ 0 & I & \mathbf{D} & \mathbf{b} \end{bmatrix}$$

### 1.3.5 Introduction to Simplex Algorithm/Method

Let a linear program be given by a canonical tableau. The simplex algorithm proceeds by performing successive pivot operations with each giving an improved basic feasible solution; the choice of pivot element at each step is largely determined by the requirement that this pivot improves the solution.

For example, consider the production function: Max Z= 200x1 +240x2

Subject to

30x1 + 15x2 ≤ 2400

20x1 + 30x2 ≤ 2400

First thing to do is to get rid of (≤) by introducing the slack element.
Thus, we rewrite the model as:

30x1 + 15x2 + s1 + 0s2 = 2400 (1)

20x1 + 30x2 + 0s1 + s2 =2400 (2)

The above equation fills the stage one of table 5.1

**Table 5.1: Simplex Algorithm/Method**

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| **Stage 1** | | | | | | | |
| Pj | - | - | 200 | 240 | 0 | 0 | Ө |
| | Activity | Resources | X1 | X2 | S1 | S2 | - |
| 0 | S1 | 2400 | 30 | **15** | 1 | 0 | 160 |
| 0 | **S2** | **2400** | **20** | **[30]** | **0** | **1** | **80** |
| | | | | | | | |
| | Zj | 0 | 0 | **0** | 0 | 0 | |
| | Pj - Zj | - | 200 | **240** | 0 | 0 | |
| **Stage 2** | | | | | | | |
| 0 | **S1** | **1200** | **19.95** | **0** | **1** | **-0.45** | **60.15** |
| 240 | X2 | 80 | **0.67** | 1 | 0 | 0.03 | 119.4 |
| | | | | | | | |
| | Zj | 19,200 | **160.8** | 240 | 0 | 7.2 | |
| | Pj - Zj | - | **39.2** | 0 | 0 | -7.2 | |
| **Stage 3** | | | | | | | |
| 200 | X1 | 60.2 | 1 | 0 | 0.05 | -0.02 | |
| 240 | X2 | 39.7 | 0 | 1 | -0.03 | 0.04 | |
| | Zj | 21568 | 200 | 240 | 2.8 | 70.4 | |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| | Pj - Zj | - | 0 | 0 | -2.8 | -70.4 | |

Note: stage three (3) maximises the equation.

The bolded row and column are both our pivot row and column, while the pivot element is acquired by observing the highest value in row Pj – Zj and the lowest values in column Ө. The pivot value is the value is block bracket ([30]) in stage 1.

For plan 2, we make x2 the subject of the formula in Equation 2.

20x1 + 30x2 + 0s1 + s2 =2400
30x2 = 2400 - 20x1 - 0s1 - s2

Dividing through by 30 gives:

x2 = 80 - 0.67x1 - 0s1 - 0.03s2 (3)

Imputing equation (4) into (1) gives:

30x1 + 15(80 - 0.67x1 - 0s1 - 0.03s2) + s1 + 0s2 = 2400
30x1 + 1200 – 10.05x1 – 0s1 – 0.45s2 + s1 + 0s2 = 2400
19.95x1 + s1 – 0.45s2 = 1200 (4)

Making x1 the subject of the formula:

x1 = 60.2 – 0.05s1 + 0.02s2 (5)

Imputing (5) into (3) gives:

x2 = 80 - 0.67x1 - 0s1 - 0.03s2

Rearranging the above equation gives:

0.67x1 + 1x2 -0s1 + 0.03s2 = 80

Thus, our model gives:

0.67(60.2 – 0.05s1 + 0.02s2) + x2 + 0.03s2 = 80
40.3 – 0.03s1 + 0.01s2 + x2 + 0.03s2 = 80 x2 – 0.03s1 +0.04s2 = 39.7

The maximizing equation is thus:

200x1 + 240x2

200 (60.2) + 240 (39.7) = 21,568.

This maximizes the value Z. i.e., revenue is maximized at the production of 60.2x1 and 39.7x2 goods.

## 1.4 Conclusion

From our discussion so far on linear programming, we can infer the following:

First, the problem of linear programming is that of choosing non- negative values of certain variables so as to maximise or minimise a given linear function subject to a given set of linear inequality constraints.

Second, the objective function according to the linear programming assumption is that the objective function is linear and that there is an additivity of resources and activities.

Third, the canonical tableau is presented as a better way of presenting linear programming in standard form.

Fourth, simplex algorithm performs successive pivot operations which each gives an improved basic feasible solution, and the choice of pivot element at each step is largely determined by the requirement that the pivot improve the solution.

## 1.5 Summary

The concept of linear programming problem is that of choosing nonnegative values of certain variables so as to maximise or minimise a given linear function, subject to a given set of linear inequality constraints.

In order to construct the model, the transformation of a linear program to one in standard form may be required.

The standard form of presenting a linear problem was discussed.

Similarly, the simplex algorithm method of analysing linear programming was discussed.

## 1.6 Tutor-Marked Assignment

1. max Z = 4x1 + 6x2

Subject to: $x1 + 2x2 \leq 20$

$3x1 + 2x2 \leq 30$

$4x1 + x2 \leq 32.5$

2. What are the assumptions of Linear Programming?
3. What are the advantages and limitation of Linear Programming?
4. max $Z = 4x1 + 6x2 + 8x3$

   Subject to: $2x1 + x2 + 3x3 \leq 80$ $x1 + 5x2 + 2x3 \leq 100$

   $3x1 + 2x2 + x3 \leq 60$

## 1.7 References/Further Reading

Micheal, D. I. (2002). "Mathematical Optimisation and Economic Theory." *An applied Mathematics.* Philadelphia: Society for Industrial and Applied Mathematics.

Jan, S. (2005). *Practical Mathematical Optimization: An Introduction to Basic Optimization Theory and Classical and New Gradient- based Algorithms (Applied Optimization).* Springer.

Mokhtar, *et al*. (2010). *Linear Programming and Network Flows.* (4th ed.). Wiley-Blackwell.

Jiri, M . (2013). *Understanding and Using Linear Programming*, Springer.

## 1.8 Answers to Self-Assessment Exercises

**Question**

What are the advantages and disadvantages of linear programming?

**Model Answer**

Advantages

v. It helps decision-makers to use their productive resources effectively.

vi. The decision-making approach of the user becomes more objective and less subjective.

vii. In a production process, bottlenecks may occur. For example, in a factory, some machines may be in great demand, while others may lie idle for some times.

viii. A significant advantage of linear programming is highlighting such bottle necks.

Disadvantages

i, Linear programming is applicable only to problems where the constraints and objective function are linear i.e., where they can be expressed as equations which represents straight lines. In real life situations, when constraints or objective functions are not linear, this technique cannot be used.

ii. Factors such as uncertainty and time are not taken into consideration.

iii. Parameters in the model are assumed to be constant but in real life situations, they are not constant.

iv. Linear programming deals with only single objectives, whereas in real life situations, we may have multiple and conflicting objectives.

v. In solving linear programming problem, there is no guarantee that we will get an integer value. In some cases of no men/machine, a non-integer value is meaningless.

# Unit 2 General Equilibrium

**Contents**



## 2.1 Introduction

There were many interesting and important contributions to the science of economics during the neo-classical era. Two of the most notable writers whose ideas have become important to the subsequent development of economic analysis were Alfred Marshall and Neon Walras. Both men were concerned with the theoretical foundation of price formation. Although each man approached the subject of market equilibrium from a different theoretical stance, yet each had a resounding impact on the science of economics. While Marshall developed the framework for partial equilibrium, Walras developed the contemporary framework for general equilibrium analysis.

General equilibrium theory as a concept of theoretical economics seeks to explain the behaviour of supply, demand, and prices in an economy with several or many interacting markets, by seeking to prove that a set of prices exists that will result in an overall equilibrium. General equilibrium theory both study economies using the model of equilibrium pricing and seeks to determine the circumstances in which the assumptions of general equilibrium will hold.

## 2.2 Objectives

At the end of this unit, you should be able to:

- differentiate between partial and general equilibrium
- state the assumptions of Walras theory
- analyse the household and consumer equilibrium
- discuss the incidence in two markets coming together to form general equilibrium.

## 2.3 Main Content

### 2.3.1 Assumptions of the Walrasian System

1. In the Walrasian system, the price which equilibrates the supply and demand sides of the market for consumer goods were assumed to be established by an auctioning process. No transaction took place until equilibrium was established in the same manner.
2. He assumed the existence of perfect competition in the commodities and factor markets. This assumption is central to general equilibrium theory because it is only within perfectly competitive system that a general equilibrium solution exists.
3. For the purpose of simplicity, he assumed the existence of only two factors of production, labour and capital (L and K). They are assumed to be homogenous and perfectly divisible.
4. For the purpose of simplicity, he assumed the existence of only two individuals A and B producing and exchanging two commodities X and Y. Their preferences for the two commodities are given by their respective indifference curves. Given the state of technology, the properties of the production functions are given by their respective isoquant maps.

5. The optimizing goal of the consumers is utility maximization subject to the constraints of their incomes and the prices of the commodities.
6. The optimizing goal of each firm is profit maximization subject to the constraint of their budgets
7. The factors of production are completely owned by the households who constitute the ultimate consumers.
8. There is full employment of all factors of production and their owners spend all incomes received.

In this system, general equilibrium is attained when the four markets (2 commodities (X and Y) and 2 factor markets (L and K) are cleared at a set of prices Px, Py, w and r respectively, and the participating institutions namely consumers and producers are in equilibrium. A general equilibrium solution involves the determination of the values of the following variables:

(i) The total quantities of commodities x and y to be produced by the firms.

(ii) The quantities of L and K to be employed in their production.

(iii) The distribution of the two commodities between the two individuals A and B.

(iv) The equilibrium prices Px, Py, w and r

**Self-Assessment Exercise**

List the Walrasian assumption of general equilibrium

## 2.3.2 The Household and Consumer Equilibrium

Households perform dual functions in general equilibrium analysis. They constitute the ultimate consumer of output. They are also the suppliers of factors of production to the business firms for the production of goods and services, which they turn around and purchase. With this, we can observe a circular interdependence subsisting between the two sectors. Households maximize their utility in consumption by equating the ratio of marginal utilities with the price ratios of the products they consume. We can express this criterion quantitatively for the two products as:

$$\frac{MUa}{MUb} = \frac{Pa}{Pb} \qquad (1)$$

Recall that the ratio of marginal utilities measures the slope of indifference curves otherwise known as the MRS. Thus, given two commodities X and Y, equilibrium is attained when:

$$MRSxy = \frac{Px}{Py} \qquad (2)$$

In general equilibrium, the MRS for the two commodities is the same for all consumers regardless of their preferences, age, sex or wealth. Households sell productive factors to the business firms. They maximise their utility by selling their factors to the highest bidder, and equilibrium in the factor market is attained in the same manner as above.

**Consumers Equilibrium**

Recall that in our analysis, we have assumed a hypothetical economy of two individuals (A and B) consuming and exchanging two commodities X and Y. From equation 2 above, the condition for joint consumer's equilibrium is given as:

$$MRS'Axy = MRS'Bxy = \frac{Px}{Py} \qquad (3)$$

Equation 3 states that MRSxy for the two consumers are identical and must equal the price ratios of the products.

**Business Firms and Production Equilibrium**

Business firms employ productive resources in the production of goods and services. The economic decisions of the firms involve the determination of what to produce, how to produce, what quantities to produce and what quantities of required factors to employ. The quantities to produce depend on the prices of the product and cost of production which in turn depends on the quantities of required resources and their respective prices. The condition for equilibrium for the firm is given as follows:

$$MRSlk = \frac{w}{r} \quad (4)$$

And MR = MC (5)

Where w = wage rate and r = cost of capital.

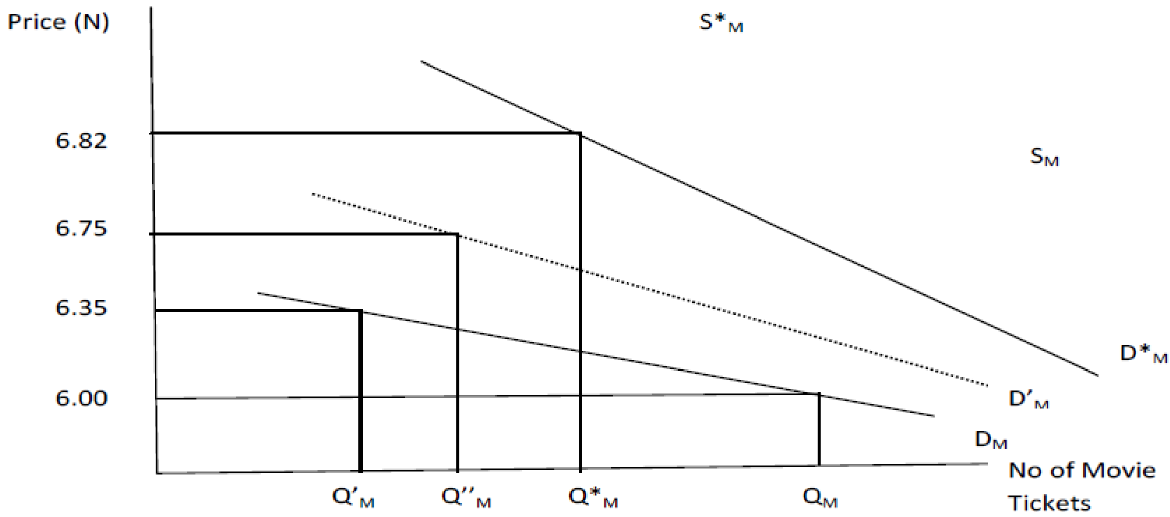
**Self-Assessment Exercise**

How is equilibrium attained in the household and firm?

### 2.3.3 Two Interdependent Markets Moving to G-E

To study the interdependence of markets, let us examine the competitive markets for DVD rentals and movie theater tickets. The two markets are closely related because DVD players give most consumers the option of watching movies at home as well as at the theater. Changes in pricing policies that affect one market are likely to affect the other, which will in turn cause feedback effects in the first market. Figure 5.1 shows the supply and demand curve for DVD's and movies. In graph (a), the price of movie tickets is initially ₦6; the market is in equilibrium at the intersection of $D_M$ and $S_M$. In part (b), the DVD market is also in equilibrium with a price of ₦3.

Now suppose that the government places a tax of ₦1 on each movie ticket purchased. The effect of this tax is determined on a partial equilibrium basis by shifting the supply curve for movies upward by ₦1, from $S_M$ to $S_M$* in graph (a) of figure 5.1. Initially, this shift causes the price of movie to increase to ₦6.35 and the quantity of movie tickets sold to fall from $Q_M$ to $Q'_M$. This is as far as partial equilibrium analysis takes us. We can go further with a general equilibrium analysis by doing two things: (1) looking at the effects of the movie tax on the market for DVD's and (2) seeing whether there are any feedback effects from the DVD market to the movie market.

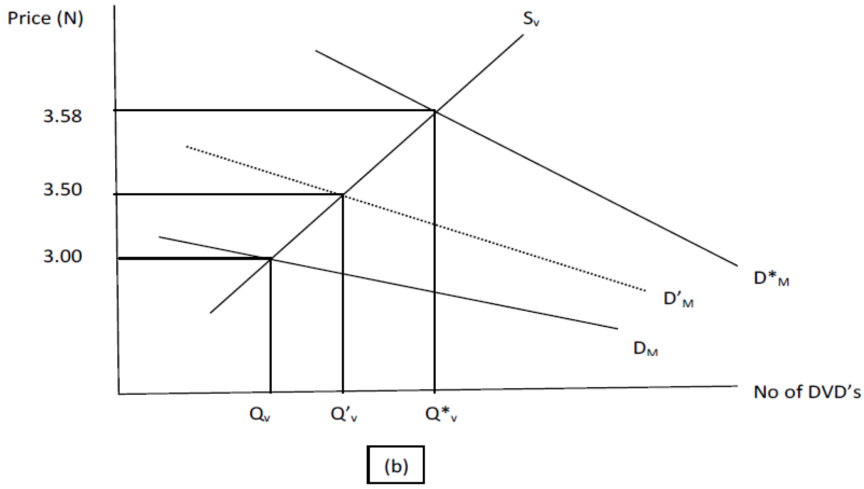



(a)

The movie tax affects the market for DVD's because movies and DVD's are substitutes. A higher movie price shifts the demand for DVD's from $D_V$ and $D'_V$ in graph (b) of figure 5.1. In turn, this shift causes the rental price of DVD's to increase from ₦3 to ₦3.5.

The original demand for movies presumed that the price of DVD was unchanged at ₦3. But because that price is now ₦3.5, the demand for movies will shift upward, from $D_M$ to $D'_M$ in graph (a) of figure 5.1. The new equilibrium price of movies is ₦6.75, instead of ₦6.35, and the quantity of movie tickets purchased has increased from $Q'_M$ to $Q''_M$. Thus, a partial equilibrium analysis would have underestimated the effect of the tax on the price of movies. The DVD market is so closely related to the market for movies that to determine the tax's full effect, we need a general equilibrium analysis.

**Fig. 5.1: Supply and Demand curves for (a) Movie Tickets and (b)DVD's**




## Reaching General Equilibrium

Our analysis is not yet complete. The change in the market price of movies will generate a feedback effect on the price of DVDs that, in turn, will affect the price of movies, and so on. In the end, we must determine the equilibrium prices and quantities of both movies and DVDs simultaneously. The equilibrium movie price of ₦6.82 is given in graph (a) of figure 5.1 by the intersection of the equilibrium supply and demand curves for movie tickets ($S_M$* and $D_M$*). The equilibrium DVD price of N3.58 is given in graph (b) of figure 5.1 by the intersection of the supply and demand curve for DVDs ($S_V$ and $D_V$*). These are the correct general equilibrium prices because the DVD market supply and demand curves have been drawn on the assumption that the price of movie tickets is ₦6.82. Likewise, the movie ticket curves have been drawn on the assumption that the price of DVDs is ₦3.58. In other words, both sets of curves are consistent with the prices in related markets, and we have no reason to expect that the supply and demand curves in either market will shift further. To find the general equilibrium prices in practice, we must simultaneously find two prices that equate quantity demanded and quantity supplied in all related markets. For our two markets, we need to find the solution to four equations (supply of movie tickets, demand for movie tickets, supply of DVDs and demand for DVDs).

Note that even if we were only interested in the market for movies, it would be important to account for the DVD market when determining the impact of a movie tax. In this example, partial equilibrium analysis would lead us to conclude that the tax will increase the price of movie tickets from ₦6 to ₦6.35. A general equilibrium analysis, however, shows us that the impact of the tax on the price of movie tickets is greater: it would in fact increase to ₦6.82.

Movies and DVD's are substitute goods. By drawing figures analogous to those in figure 5.1, you should be able to convince yourself that if the goods in question are complements, a partial equilibrium analysis will overstate the impact of a tax. Think about petroleum and automobiles, for example. A tax on petroleum will cause its price to go up, but this increase will reduce demand for automobiles, which in turn reduces the demand for petroleum, causing its price to fall somewhat.

**Self-Assessment Exercise**

How is general equilibrium attained in two different markets?

## 2.3.4 Efficiency in Exchange and the Advantage of Trade

To examine the concept of economic efficiency, we start with an exchange economy, analysing the behaviour of two consumers who can trade either of two goods between themselves. This analysis also applies to trade between two countries. Suppose the two goods are initially allocated so that both consumers can make themselves better off by trading with each other. In this case, the initial allocation of goods is economically inefficient. In an efficient allocation of goods, no one can be made better off without making someone else worse off.

**Advantage of Trade**

As a rule, voluntary trade between two people or two countries is mutually beneficial. To see how trade makes people better off let us consider a two-person exchange, assuming that the exchange itself is costless. Suppose Tola and Bola have

10 units of food and 6 units of clothing between them. Table 5.2 shows that initially, Tola has 7 units of food and 1 unit of clothing, and Bola has 3 units of food and 5 units of clothing. To decide whether trade would be advantageous, we need to know their preferences for food and clothing. Suppose that because Bola has a lot of clothing and little food, her marginal rate of substitution (MRS) for food is 3: to get 1 unit of food, she will give up 3 units of clothing. However, Tola's MRS of food for clothing is only ½: he will give up only ½ unit of clothing to get 1 unit of food.

**Table 5.2: Marginal Rate of Substitution**

| Individual | Initial Allocation | Trade | Final Allocation |
|---|---|---|---|
| Tola | 7F, 1C | -1F, +1C | 6F, 2C |
| Bola | 3F, 5C | +1F, -1C | 4F, 4C |

There is thus room for mutually advantageous trade because Tola values clothing more highly than Bola does, whereas, Bola values food more highly than Tola does. To get another unit of food, Bola would be willing to trade up to 3 units of clothing. But Tola will give up 1 unit of food for ½ unit of clothing. The actual terms of trade depend on the bargaining process. Among the possible outcomes are trade of 1unit of food by Tola for anywhere between ½ and 3 units of clothing from Bola.

Suppose Bola offers Tola 1 unit of clothing for 1 unit of food, and Tola agrees, both will be better off. Tola will have more clothing, which he values more than food. Whenever two consumer's MRS's are different, there is room for mutual trade because the allocation of resources is inefficient. Through trading, both consumers will be made better off.

**Self-Assessment Exercise**

Write an essay on how efficient trade can be carried out between two consumers with different initial allocations.

### 2.3.5 Theory of Wage Distribution

In analyzing the earnings of labour, it is necessary to take into account the imperfections of the labour market and the actions of trade unions. Imperfections in the market make for a certain amount of indeterminacy in which considerations of fairness, equity, and tradition play a part. These affect the structure of wages - i.e., the relationships between wages for various kinds of labour and various skills. Therefore, one cannot say that the income difference between a carpenter and a physician, or between a bank clerk and a truck driver, is completely determined by marginal productivity, although it is true that in the long run the wage structure is influenced by supply and demand. The role of the trade unions has been a subject of much debate. The naive view that unions can raise wages by their efforts irrespective of market forces is, of course, incorrect. In any particular industry, exaggerated wage claims may lead to a loss of employment; this is generally recognized by union leaders. The opposite view, that trade unions cannot influence wages at all (unless they alter the basic relationship between supply and demand for labour), is held by a number of economists with respect to the real wage level of the economy as a whole. They agree that unions may push up the money wage level, especially in a tight labour market, but argue that this will lead to higher prices and so the real wage rate for the economy as a whole will not be increased accordingly. These economists also point out that high wages tend to encourage substitution of capital for labour (the cornerstone of neoclassical theory). These factors do indeed operate to check the power of trade unions, although the extreme position that the unions have no power at all against the iron laws of the market system is untenable. It is safe to say that basic economic forces do far more to determine labour's share than do the policies of the unions. The main function of the unions lies rather in modifying the wage structure; they are able to raise the bargaining power of weak groups of workers and prevent them from lagging behind the others.

**Interest and Profit**

The earnings of capital are determined by various factors. Capital stems from two sources: from saving (by households, financial institutions, and businesses) and from the creation of money by the banks. The creation of money depresses the rate of interest below what may be called its natural rate. At this lower rate, businessmen will invest more, the capital stock will increase, and the marginal productivity of capital will decline. Although this chain of reactions has drawn the attention of monetary theorists, its impact on income distribution is probably not very important, at least not in the long run. There are also other factors, such as government borrowing, that may affect the distribution of income; it is difficult to say in what direction. The basic and predominant determinant is marginal productivity; the continuous accumulation of capital depresses the rate of interest.

One type of earning that is not explained by the neoclassical theory of distribution is profit, a circumstance that is especially awkward because profits form a substantial part of national income (20–25 per cent); they are an important incentive to production and risk taking as well as being an important source of funds for investment. The reason for the failure to explain profit lies in the essentially static character of the neoclassical theory and in its preoccupation with perfect competition. Under such assumptions, profit tends to disappear. In the real world, which is not static and where competition does not conform to the theoretical assumptions, profit may be explained by five causes. One is uncertainty; an essential characteristic of business enterprise is that not all future developments can be foreseen or insured against. Frank H. Knight (1921) introduced the distinction between risk, which can be insured for and thus treated as a regular cost of production, and uncertainty, which cannot. In a free enterprise economy, the willingness to cope with the uninsurable has to be remunerated, and thus it is a factor of production. A second way of accounting for profits is to explain them as a premium for introducing new technology or for producing more efficiently than one's competitors. This dynamic element in profits was stressed by Joseph Schumpeter (1911). In his view, prices are determined by the level of costs in the least progressive firms; the firms that introduce a new product or a new method will benefit from lower costs than its competitors. A third source of profits is monopoly and related forms of market power, whether deliberate as with cartels and other restrictive practices or arising from the industrial structure itself. Some economists have developed theories in which the main influence determining distributive shares is the relative “degree of monopoly” exerted by various factors of production, but this seems a bit one-sided. A fourth source of profits is sudden shifts in demand for a given product-so-called windfall profits, which may be accompanied by losses elsewhere. Finally, there are profits arising from general increases in total demand caused by a certain kind of inflationary process when costs, especially wages, lag behind rising prices. Such is not always the case in modern inflations.

**Rent**

In classical economics, rent was the return to an "owner" of land. In later economic theory, this term is expanded as economic rent to include other forms of unearned income typically realized from barriers to entry. Land ownership is considered to be a barrier to entry because land owners make no contribution to the production process. They simply prevent others from using that which would otherwise be useful.

Rent is that portion of the produce of the earth, which is paid to the landlord for the use of the original and indestructible powers of the soil. It is often, however, confounded with the interest and profit of capital, and, in popular language, the term is applied to whatever is annually paid by a farmer to his landlord.

**Self-Assessment Exercise**

Discuss the theory of distribution

## 2.4 Conclusion

The concept of general equilibrium focuses on the explanation of the supply, demand and prices in the whole economy with several interacting markets. Households and consumers attain equilibrium when the marginal utility of a particular good is equal to its market price, while for firms, it is when the marginal revenue is equal to its marginal cost. Mutual trade between consumers exists so as to make them better- off than they were before.

## 2.5 SUMMARY

Partial equilibrium analysis of markets assumes that related markets are unaffected. General equilibrium analysis examines all markets simultaneously, taking into account feedback effects to other markets on the market being studied.

An allocation is efficient when no consumer can be made better off by trade without making someone else worse off. When consumers make all mutually advantageous trades, the outcome is Pareto efficient.

A competitive equilibrium describes a set of prices and quantities. When each consumer chooses her most preferred allocation, the quantity demanded is equal to the quantity supplied in every market.

Efficiency in the allocation of goods to consumers is achieved only when the marginal rate of substitution of one good for another in consumption is equal to the marginal rate of transformation of one good for another in production.

Free international trade expands a country's production possibilities frontier, and as a result, consumers are better off.

## 2.6 Tutor-Marked Assignment

1. Discuss the theory of Distribution.
2. Discuss the household and firm equilibrium.
3. Discuss the advantages of trade.
4. Explain the assumptions of the Walrasian system of general equilibrium.

## 2.7 References/Further Reading

Colander, D. C. (2013). *Economics*. (9th ed.). New York: McGraw-Hill.

Cowen, T. & Alex, T. (2012). *Modern Principles: Economics*. (2nd ed.). New York: Worth Publishers.

Koutsoyiannis, A. (1987). *Modern Microeconomics*. (2nd ed.).

Robert, S. P. & Daniel, L. R. (2009). *Microeconomics.* (7th ed.). New Jersey: Pearson Education International.

## 2.8 Answers to Self-Assessment Exercises

**Question**

How is equilibrium attained in the household and firm?

**Model Answer**

Equilibrium in Household

Households perform dual functions in general equilibrium analysis. They constitute the ultimate consumer of output. They are also the suppliers of factors of production to the business firms for the production of goods and services, which they turn around and purchase. With this, we can observe a circular interdependence subsisting between the two sectors. Households maximize their utility in consumption by equating the ratio of marginal utilities with the price ratios of the products they consume. We can express this criterion quantitatively for the two products as:

$$\frac{MUa}{MUb} = \frac{Pa}{Pb} \qquad \text{------------------------------------1}$$

Recall that the ratio of marginal utilities measures the slope of indifference curves otherwise known as the MRS. Thus, given two commodities X and Y, equilibrium is attained when:

$$MRSxy = \frac{Px}{Py} \qquad \text{------------------------------------2}$$

In general equilibrium, the MRS for the two commodities is the same for all consumers regardless of their preferences, age, sex or wealth. Households sell productive factors to the business firms. They maximise their utility by selling their factors to the highest bidder, and equilibrium in the factor market is attained in the same manner as above.

Equilibrium in Firm

Business firms employ productive resources in the production of goods and services. The economic decisions of the firms involve the determination of what to produce, how to produce, what quantities to produce and what quantities of required factors to employ. The quantities to produce depend on the prices of the product and cost of production which in turn depends on the quantities of required resources and their respective prices. The condition for equilibrium for the firm is given as follows:

$MRSlk = \frac{w}{r}$ ---------------------------------4

And MR = MC ------------------------------5

Where w = wage rate and r = cost of capital.

**Question**

Write an essay on how efficient trade can be carried out between two consumers with different initial allocations.

**Model Answer**

As a rule, voluntary trade between two people or two countries is mutually beneficial. To see how trade makes people better off let us consider a two-person exchange, assuming that the exchange itself is costless. Suppose Tola and Bola have 10 units of food and 6 units of clothing between them. The table below shows that initially, Tola has 7 units of food and 1 unit of clothing, and Bola has 3 units of food and 5 units of clothing. To decide whether trade would be advantageous, we need to know their preferences for food and clothing. Suppose that because Bola has a lot of clothing and little food, her marginal rate of substitution (MRS) for food is 3: to get 1 unit of food, she will give up 3 units of clothing. However, Tola's MRS of food for clothing is only ½: he will give up only ½ unit of clothing to get 1 unit of food.

**Marginal Rate of Substitution**

| Individual | Initial Allocation | Trade | Final Allocation |
|---|---|---|---|
| Tola | 7F, 1C | -1F, +1C | 6F, 2C |
| Bola | 3F, 5C | +1F, -1C | 4F, 4C |

There is thus room for mutually advantageous trade because Tola values clothing more highly than Bola does, whereas, Bola values food more highly than Tola does.

To get another unit of food, Bola would be willing to trade up to 3 units of clothing. But Tola will give up 1 unit of food for ½ unit of clothing. The actual terms of trade depend on the bargaining process. Among the possible outcomes are trade of 1unit of food by Tola for anywhere between ½ and 3 units of clothing from Bola.

Suppose Bola offers Tola 1 unit of clothing for 1 unit of food, and Tola agrees, both will be better off. Tola will have more clothing, which he values more than food. Whenever two consumer's MRS's are different, there is room for mutual trade because the allocation of resources is inefficient. Through trading, both consumers will be made better off.

## Unit 3 Social Welfare

**Contents**



### 3.1 Introduction

Welfare economics is concerned with the evaluation of alternative economic situations from the point of view of the society's well-being. Assume that the total welfare in a country is W, but given the factor endowment (resources) and the state of technology, suppose that this welfare could be larger, for example, W*. The tasks of welfare economics are to (1) show that in the present state, $W < W^*$, and (2) to seeks ways of raising W to W*.

### 3.2 Objectives

At the end of this unit, you should be able to:

- describe the different criteria of social welfare
- explain the social welfare function
- apply the individualist approach to welfare maximization
- determine a fair and equal allocation in the society.

## 3.3 Main Content

### 3.3.1 Criteria of Social Welfare

To evaluate alternative economic situations, we need some criterion of social wellbeing or welfare. The measurement of social welfare requires some ethical standard and interpersonal comparison, both of which involve subjective value judgments. Objective comparisons and judgments of the deservingness or worthiness of different individuals are virtually impossible. Various criteria of social welfare have been suggested by economists at different times.

**a. Growth of GNP Criterion**

Adam Smith implicitly accepted the growth of the wealth of a society, that is, the growth of the gross national product, as a welfare criterion. He believed that economic growth resulted in the increase of social welfare because growth increased employment and the goods available for consumption to the community. To Adam Smith, economic growth meant bringing W closer W*.

The growth criterion implies acceptance of the status quo of income distribution as ethical or just. Furthermore, growth may lead to a reduction in the social welfare, depending on who avails mostly from it. However, the growth criterion highlights the importance of efficiency in social welfare, given that social welfare depends on the amount of goods and services as well as its means of distribution.

**b. Bentham's Criterion**

Jeremy Bentham, an English economist, argued that welfare is improved when the greatest good is secured for the greatest number. Implicit in this dictum is the assumption that the total welfare is the sum of the utilities of the individuals of the society. The application of the Bentham's ethical system to economics has serious shortcomings. To highlight the pitfall in this criterion, let us assume that the society consists of three individuals, A, B and C, so that

$W = U_A + U_B + U_C$

According to Bentham, $\Delta W > 0$ if $(\Delta U_A + \Delta U_B + \Delta U_C) > 0$. However, assume that the change which resulted in the changes in the individuals' utilities is such that A's and B's utility increases, while C's utility decreases, but $(\Delta U_A + \Delta U_B) > \Delta U_C$. In other words, two individuals are better-off while the third is worse-off after the change has taken place, but the sum of the increase in utilities of A and B is greater that the decrease in the utility of C. This criterion implies that A and B have greater worthiness than C. That is, implicit in Bentham's criterion is an interpersonal comparison of the deservingness of the members of the society. Another difficulty of this criterion is that it cannot be applied to compare situations where the greatest good and the greatest numbers do not exist simultaneously.

**c. Cardinalist Criterion**

Several economists proposed the use of the law of diminishing marginal utility as a criterion of welfare. To illustrate their argument, assume that the society consists of three individuals: A has an income of 1000 naira, while B and C have income of 500 naira each. Consumer A can buy double quantities of goods compared to B and C. However, given the law of diminishing marginal, A's total utility is less than double the total utility of either B or C; because A's marginal utility of money is less than that of B and C. Thus, $W < W^*$. To increase social welfare, income should be redistributed among the three individuals. In fact, cardinal welfare maintained that social welfare would be maximized if income was equally distributed to all members of the society. The cardinalist approach to welfare has a serious flaw; it assumes that all individuals have identical utility functions for money, so that with an equal income distribution, all would have the same marginal utility of money. This assumption is too strong. Individuals differ in their attitude towards money. A rich person may have a utility for money function that lies far above the utility function of poorer individuals. In this case, a redistribution of income might reduce total welfare.

### d. Pareto Optimality Criterion

The criterion refers to economic efficiency which can be objectively measured. It is called the Pareto criterion after the famous Italian economist Vilfredo Pareto. According to this criterion, any change that makes at least one individual better-off and no one worse-off is an improvement in social welfare. Conversely, a change that makes no one better-off and at least one worse-off is a decrease in social welfare. Stating this criterion differently, a situation in which it is impossible to make anyone better-off without making someone worse-off is said to be Pareto-optimal or Pareto-efficient. The three marginal conditions accepted for the attainment of a Pareto-efficient situation in an economy is stated below:

1. Efficiency of distribution of commodities among consumers (efficiency in exchange)
2. Efficiency of the allocation of factors among firms (efficiency of production)
3. Efficiency in the allocation of factors among commodities (efficiency in product-mix, or composition of output).

### e. The Kaldor-Hicks Compensation Criterion

Nicholas Kaldor and John Hicks suggested the following approach to estimating a welfare criterion. Assume that a change in the economy is being considered, which will benefit some (gainers) and hurt some (losers). One can ask the gainers how much money they would be prepared to pay in order to have the change and the losers how much money they would be prepared to pay in order to prevent the change. If the amount of money of the gainers is greater than the amount of the losers, the change constitutes an improvement in welfare, because the gainers could compensate the losers and still have some net gain. Thus, the Kaldor-Hicks compensation criterion states that a change constitutes an improvement in social welfare if those who benefit from it could compensate those who are hurt, and still be left with some net gain.

The Kaldor-Hicks criterion evaluates alternative situation on the basis of monetary evaluation of different persons. Thus, it implicitly assumes that the marginal utility of money is the same for all the individuals in the society. Given that income distribution is unequal in the real world, this assumption is absurd. Assume, for example, that the economy consists of two individuals, A, who is a millionaire, and B, who has an income of N4000. Suppose that the change (being considered by the government) will benefit A, who is willing to pay ₦2000 for this change to happen, while it hurts B, who is prepared to pay ₦1000 to prevent the change. According to the Kaldor-Hicks criterion, the change will increase the social welfare (since the net gain to A, after he compensates B, is ₦1000). However, the gain of ₦2000 gives very little additional utility to millionaire A, while the loss of ₦1000 will decrease a lot the well-being of B who has greater marginal utility of money than A. Thus, the total welfare will be reduced if the change takes place. Only if the marginal utility of money is equal for all the individuals would the Kaldor-Hicks criterion be a correct welfare measure.

**Self-Assessment Exercise**
Identify the different criteria of social welfare

### 3.3.2 Social Welfare Maximization

The conditions for social welfare maximization in the simple two factors, two commodities, and two consumer model are stated below.

1. There are two factors, labour L and capital K, whose quantities are given (in perfectly inelastic supply). These factors are homogenous and perfectly divisible.
2. Two products, X and Y, are produced by two firms. Each firm produces only one commodity. The production function gives rise to smooth isoquants, convex to the origin, with constant returns to scale. Indivisibilities in the production process are ruled out.

3. There are two consumers whose preferences are presented by indifference curves, which are continuous, convex to the origin and do not intercept.
4. The goal of consumers is utility maximization and the goal of firms is profit maximization.
5. The production functions are independent. This rules-out joint products and external economies and diseconomies in production.
6. The utilities of consumers are independent, Bandwagon, snob and Veblen effects are ruled out. There are no external economies or diseconomies in consumption.
7. The ownership of factors, that is, the distribution of the given L and K between the two consumers, is exogenously determined.
8. A social welfare function, $W = f(U_A, U_B)$, exists. This permits a unique preference-ordering of all possible states, based on the position of the two consumers in their own preference maps. This welfare function incorporates an ethical valuation of the relative deservingness or worthiness of the two consumers.

**Self-Assessment Exercise**
What are the conditions of social welfare maximization?

### 3.3.3 Individualistic Social Welfare Functions

Up until now we have been thinking of individual preferences as being defined over entire allocations rather than over each individual's bundle of goods. As we remarked earlier however, individuals might only care about their own bundles. In this case, we can use $x_i$ to denote individual *i*'s consumption bundle, and let $u_i(x_i)$ be individual *i*'s utility level using some fixed representation of utility. Then a social welfare function will have the form:

$$W = W(u_i(x_i), \ldots, u_n(x_n)).$$

The welfare function is directly a function of the individuals' utility levels, but it is indirectly a function of the individual agents' consumption bundles. This special form of welfare function is known as an individualistic welfare function or a Bergson-Samuelson welfare function. If each agent's utility depends only on his or her own consumption, then there are no consumption externalities. Thus, we have an intimate relationship between Pareto efficient allocations and market equilibria: All competitive equilibria are Pareto efficient, and, under appropriate convexity assumptions, all Pareto efficient allocations are competitive equilibria. Now we can carry this categorization one step further. Given the relationship between Pareto efficiency and welfare maxima described above, we can conclude that all welfare maxima are competitive equilibria and that all competitive equilibria are welfare maxima for some welfare functions.

### 3.3.4 Fair Allocation

The welfare function approach is a very general way to describe social welfare. As it is so general, it can be used to summarize the properties of many kinds of moral judgments. On the other hand, it isn't much used in deciding what kinds of ethical judgments might be reasonable ones. Another approach is to start with some specific moral judgments and then examine their implications for economic distribution. This is the approach taken in the study of fair allocations. We start with a definition of what might be considered a fair way to divide a bundle of goods, and then use our understanding of economic analysis to investigate its implications. Suppose that you were given some goods to divide fairly among (n) equally deserving people. How would you do it? It is probably safe to say that in this problem most people would divide the goods equally among the n agents. Given that they are by hypothesis equally deserving, what else could you do? What is appealing about this idea of equal division? One appealing feature is that it is symmetric. Each agent has the same bundle of goods; no agent prefers any other agent's bundle of goods to his or her own, since they all have exactly the same thing.

Unfortunately, an equal division will not necessarily be Pareto efficient. If agents have different tastes, they will generally desire to trade away from equal division. Let us suppose that this trade takes place and that it moves us to a Pareto efficient allocation. The question arises: Is this Pareto efficient allocation still fair in any sense? Does trade from equal division inherent any of the symmetry of the starting points? The answer is not necessarily a yes. Consider the following example. We have three people, A, B, and C. A and B have the same tastes, and C has different tastes. We start from an equal division and suppose that A and C get together and trade. Then they will typically both be made better off. Now B, who didn't have the opportunity to trade with C, will envy A, that is, he would prefer A's bundle to his own. Even though A and B started with the same allocation, A was luckier in her trading, and this destroyed the symmetry of the original allocation. This means that arbitrary trading from an equal division will not necessarily preserve the symmetry of the starting point of equal division. We might well ask if there is any allocation that preserves this symmetry. Is there any way to get an allocation that is both Pareto efficient and equitable at the same time?

**Self-Assessment Exercise**

How is fair allocation determined?

### 3.3.5 Envy and Equity

What do we mean by "symmetric" or "equitable"? One possible set of definitions is as follows. We say an allocation is equitable if no agent prefers any other agent's bundle of goods to his or her own. If some agent i does prefer some other agent *j*'s bundle of goods, we say that *i* envies *j*. Finally, if an allocation is both equitable and Pareto efficient, we will say that it is a fair allocation. These are ways of formalizing the idea of symmetry alluded to above. An equal division allocation has the property that no agent envies any other agent-but there are many other allocations that have this same property. Consider the Edgeworth box in figure 5.2 to determine whether any allocation is equitable or not, just look at the allocation that results if the two agents swap bundles. If this swapped allocation lies "below" each agent's indifference curve through the original allocation, then the original allocation is an equitable allocation. Note also that the allocation in figure 5.2 is also Pareto efficient. Thus, it is not only equitable, in the sense that we defined the term, but it is also efficient. By our definition, it is a fair allocation.

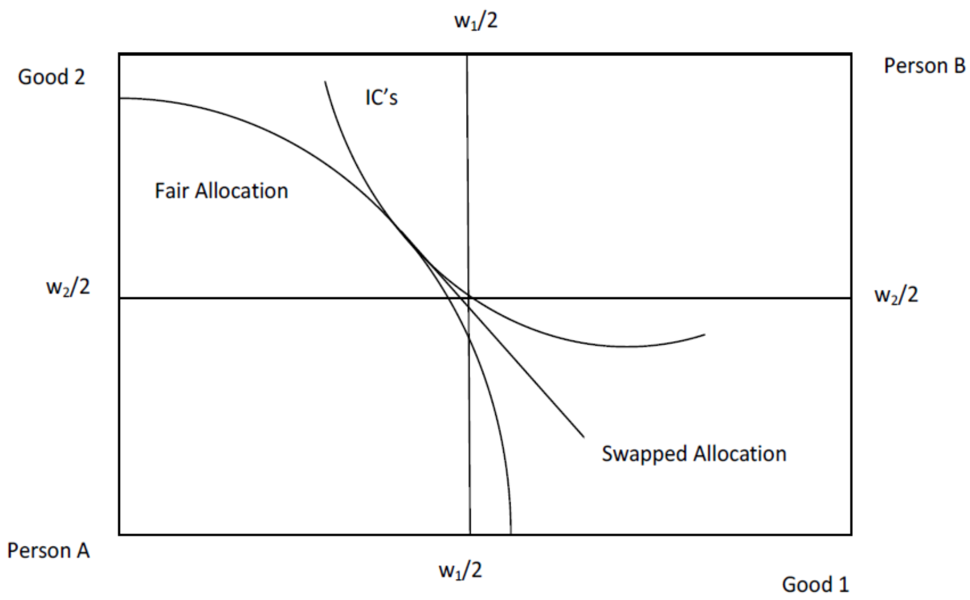



**Fig. 5.2: The Edgeworth Box**

**Self-Assessment Exercise**

How can an allocation be shared equally without envy?

## 3.4 Conclusion

From our discussion so far, we can infer that:

First, welfare functions are often used by economists of one sort or another to represent distributional judgments about allocations.

Second, as long as the welfare function is increasing in each individual's utility, a welfare maximum will be Pareto efficient. Furthermore, every Pareto efficient allocation can be thought of as maximizing some welfare function.

Third, the idea of fair allocations provides an alternative way to make distributional judgments. This idea emphasizes the idea of symmetric treatment.

Fourth, even when the initial allocation is symmetric, arbitrary methods of trade will not necessarily produce a fair allocation. However, it turns out that the market mechanism will provide a fair allocation.

## 3.5 Summary

The concept of social welfare was reviewed, with extensive discussion on different criteria of social welfare. It was established that Pareto efficient allocation and the criteria emphasizes the allocation efficiency in the society. Different conditions for social welfare maximisation were also discussed before individualistic social function was reviewed. Finally, fair allocations as well as envy and equity concept of the social welfare were discussed.

## 3.6 Tutor-Marked Assignment

1. Suppose that an allocation is Pareto efficient, and that each individual only cares about his own consumption. Prove that there must be some individual that envies no one.
2. Analyze the concept of social welfare function, fair allocation and the social welfare criterion.

## 3.7 References/Further Reading

Koutsoyiannis, A. (1987.). *Modern Microeconomics*. ($2^{nd}$ ed.).

Krugman, P. & Robin, W. (2012). *Economics*. ($3^{rd}$ ed.). New York: Worth Publishers.

Lipsey, *et al.* (2008). *Economics*. ($13^{th}$ ed.). Upper Saddle River, NJ: Pearson.

Robert, S. P. & Daniel, L. R. (2009). *Microeconomics.* ($7^{th}$ ed.). New Jersey: Pearson Education International

Jhingan, M.L. (2006). *Microeconomic Theory*. ($5^{th}$ ed.). Delhi, Vrinda Publications

Geoffrey, A.J. (2011). *Advanced Microeconomic Theory*. ($3^{rd}$ ed.), London and New York, Prentice Hall

James, B. (2005). *Microeconomic Theory*: *A Concise Course.* New York, Oxford University Press.

### 3.8 Answers to Self-Assessment Exercises

**Question**

What are the conditions of social welfare maximization?

**Model Answer**

The conditions for social welfare maximization in the simple two factors, two commodities, and two consumer model are stated below.

1. There are two factors, labour L and capital K, whose quantities are given (in perfectly inelastic supply). These factors are homogenous and perfectly divisible.
2. Two products, X and Y, are produced by two firms. Each firm produces only one commodity. The production function gives rise to smooth isoquants, convex to the origin, with constant returns to scale. Indivisibilities in the production process are ruled out.
3. There are two consumers whose preferences are presented by indifference curves, which are continuous, convex to the origin and do not intercept.
4. The goal of consumers is utility maximization and the goal of firms is profit maximization.
5. The production functions are independent. This rules-out joint products and external economies and diseconomies in production.
6. The utilities of consumers are independent, Bandwagon, snob and Veblen effects are ruled out. There are no external economies or diseconomies in consumption.
7. The ownership of factors, that is, the distribution of the given L and K between the two consumers, is exogenously determined.
8. A social welfare function, $W = f(U_A, U_B)$, exists. This permits a unique preference-ordering of all possible states, based on the position of the two consumers in their own preference maps. This welfare function incorporates an ethical valuation of the relative deservingness or worthiness of the two consumers.

**Question**

How is fair allocation determined?

**Model Answer**

The welfare function approach is a very general way to describe social welfare. As it is so general, it can be used to summarize the properties of many kinds of moral judgments. On the other hand, it isn't much used in deciding what kinds of ethical judgments might be reasonable ones. Another approach is to start with some specific moral judgments and then examine their implications for economic distribution. This is the approach taken in the study of fair allocations. For instance, if you were given some goods to divide fairly among (n) equally deserving people. How would you do it? It is probably safe to say that in this problem most people would divide the goods equally among the n agents. Given that they are by hypothesis equally deserving, what else could you do? What is appealing about this idea of equal division? One appealing feature is that it is symmetric. Each agent has the same bundle of goods; no agent prefers any other agent's bundle of goods to his or her own, since they all have exactly the same thing.

Unfortunately, an equal division will not necessarily be Pareto efficient. If agents have different tastes, they will generally desire to trade away from equal division. Let us suppose that this trade takes place and that it moves us to a Pareto efficient allocation. The question arises: Is this Pareto efficient allocation still fair in any sense? Does trade from equal division inherent in any of the symmetry of the starting points? The answer is not necessarily a yes. Consider the following example. We have three people, A, B, and C. A and B have the same tastes, and C has different tastes. We start from an equal division and suppose that A and C get together and trade. Then they will typically both be made better off. Now B, who didn't have the opportunity to trade with C, will envy A, that is, he would prefer A's bundle to his own. Even though A and B started with the same allocation, A was luckier in her trading, and this destroyed the symmetry of the original allocation. This means that arbitrary trading from an equal division will not necessarily preserve the symmetry of the starting point of equal division.

**Question**

How can an allocation be shared equally without envy?

**Model Answer**

First, there is need to understand what we mean by "symmetric" or "equitable". One possible set of definitions is as follows. We say an allocation is equitable if no agent prefers any other agent’s bundle of goods to his or her own. If some agent i does prefer some other agent *j*’s bundle of goods, we say that *i* envies *j*. Finally, if an allocation is both equitable and Pareto efficient, we will say that it is a fair allocation. These are ways of formalizing the idea of symmetry alluded to above. An equal division allocation has the property that no agent envies any other agent-but there are many other allocations that have this same property. Consider the Edgeworth box below to determine whether any allocation is equitable or not, just look at the allocation that results if the two agents swap bundles. If this swapped allocation lies "below" each agent's indifference curve through the original allocation, then the original allocation is an equitable allocation. Note also that the allocation in the Edgeworth box is also Pareto efficient. Thus, it is not only equitable, in the sense that we defined the term, but it is also efficient. By our definition, it is a fair allocation.

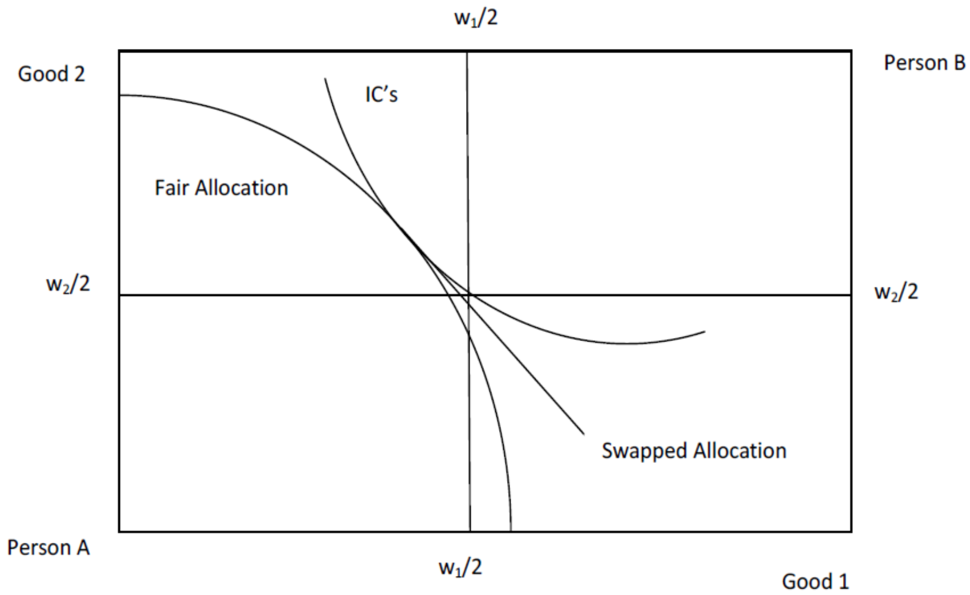



**Fig. 5.2: The Edgeworth Box**