

**NATIONAL OPEN UNIVERSITY OF NIGERIA**

**ADVANCED MICROECONOMIC THEORY  
ECO711**

**COURSE GUIDE**

**Course Developers:**

**Dr. Emmanuel Ifeanyi AJUDUA**

**Department Economics  
Faculty of Social Sciences  
National Open University of Nigeria**

**&**

**Dr. Athanasius Nnanyelugo NWOKORO**

**Department Economics  
Faculty of Social Sciences  
Imo State University, Owerri, Imo State**

**Course Editor:**

**Professor Anthony AKAMOBI**

**Department of Economics  
Chukwuemeka Odumegwu Ojukwu University  
Anambra State.**

© 2020 by NOUN Press  
National Open University of Nigeria,  
Headquarters,  
University Village,  
Plot 91, Cadastral Zone,  
Nnamdi Azikiwe Expressway,  
Jabi, Abuja.

e-mail: [centralinfo@nou.edu.ng](mailto:centralinfo@nou.edu.ng)

URL: [www.nou.edu.ng](http://www.nou.edu.ng)

All rights reserved. No part of this book may be reproduced, in any form or by any means, without permission in writing from the publisher.

Printed:

ISBN:

# **CONTENT**

Introduction

Course Content

Course Aims

Course Objectives

Working through This Course

Course Materials

Study Units

Textbooks and References

Assignment File

Presentation Schedule

Assessment

Tutor-Marked Assignment (TMAs)

Final Examination and Grading

Course Marking Scheme

Course Overview

How to Get the Most from This Course

Tutors and Tutorials

Summary

# INTRODUCTION

## Welcome to ECO711 ADVANCED MICROECONOMIC THEORY.

ECO711: Advanced Microeconomic Theory is a three-credit and one-semester Postgraduate Diploma course for students in the Department of Economics, Faculty of Social Sciences and as a service course for other Postgraduate programmes at the National Open University of Nigeria. The course will help you to gain an in-depth knowledge of the microeconomics as it has been simplified so as to accommodate postgraduate students who have little or no knowledge of economics at undergraduate level. In order to assist students', understand better, practice test and assignments would be given at the end of each unit. This course guide therefore will give you an insight to microeconomic theory in a broader way and how to make use and apply theories as an economist. It tells you about the course materials 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. Answers to your tutor marked assignments (TMAs) are therein already. The course is made up of thirteen units spread across fifteen lectures weeks. This course guide gives you an insight to Advanced Microeconomic Theory in a broader way and how to make use and apply economic theories as an economist. It tells you about the course materials 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. Answers to your tutor marked assignments (TMAs) are therein already.

### Course Content

This course is basically on Advanced Microeconomic Theory and is made up of **thirteen units (four modules)** covering areas such as fundamental quantitative relationships; General equilibrium and disequilibrium; dynamic equilibrium analysis; Theories of consumer behaviour; Production functions: duopoly, oligopoly; bilateral monopoly and monopsony. Other areas include Theories of determination of wages, rent, interest and profit; economic efficiency; and equity; externalities; social and private costs; social welfare functions; theory of the firm; linear programming and applications; theory of distribution; social welfare function; externalities.

### Course Aims

The aim is to help you have the basic knowledge of microeconomics as it relates to everyday economic activities. However, the following broad aims will also be achieved:

- To introduce you to the basic analysis and functional relationship between variables
- To expose you to the theory of consumer behavior and the analysis of utility maximization
- To expose you to equilibrium, its types and conditions for equilibrium
- To expose you extensively to market structure
- To enlighten you on the theory of distribution and production.
- To enlighten you on the concept of profit maximization

- To familiarize you with the knowledge of cost benefit analysis.
- To teach you the concept of linear programming and its application

## **Course Objectives**

Generally, the objective of ECO711 is centred on equipping postgraduate students with necessary microeconomic theories and knowledge. This will be of great use to you as an economist. Each unit in the course material has its own objective(s) which has been clearly stated at the beginning of each unit. It is advisable that you read them before working through the units. References may be made to them in the course of studying the units. On the successful completion of the course, you should be able to:

- Discuss the functional relationship amongst variables
- Understand the utility analysis
- Define concepts as it relates to utility, production and market structures
- Understand consumer behavior and the relationship between total Utility and marginal utility
- Understand partial, general dynamic, stable, unstable and neutral equilibrium
- Be conversant with types of production function, duopolies, bilateral monopoly and monopsony
- Understand Profit-Maximizing Theories
- Understand Concept of Theory Distribution
- Understand the meaning and application of Cost Benefit Analysis

## **Working Through the 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 exercises called Student Assessment Exercises (SAE). At some points in the course, you will be required to submit assignments for assessment purposes. At the end of the course there is a final examination. This course should take about 15weeks to complete and some components of the course are outlined under the course material subsection.

## **Course Material**

The major component 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 follows:

1. Course guide
2. Study unit
3. Textbook
4. Assignment file
5. Presentation schedule

## **Study Unit**

There are thirteen (13) units in this course which should be studied carefully and diligently.

## **MODULE 1: QUANTITATIVE RELATIONSHIPS, CONSUMER BEHAVIOUR AND EQUILIBRIUM ANALYSIS**

- Unit 1: Fundamental Quantitative Relationships**
- Unit 2: Theory of Consumer Behaviour**
- Unit 3: Equilibrium Analysis**

## **MODULE 2: PRODUCTION FUNCTIONS, MARKET STRUCTURES, DETERMINATION OF WAGES, RENT, INTEREST AND PROFIT**

- Unit 1: Production Function, Duopoly and Oligopoly**
- Unit 2: Bilateral Monopoly and Monopsony**
- Unit 3: Theories of Determination of Wages, Rent, Interest and Profit**

## **MODULE 3: THEORIES OF FIRM, DISTRIBUTION AND COST BENEFIT ANALYSIS**

- Unit 1: Theory of The Firm**
- Unit 2: Theory of Distribution**
- Unit 3: Cost-Benefit Analysis**

## **MODULE 4: WELFARE ECONOMICS AND LINEAR PROGRAMMING**

- Unit 1: Social Welfare**
- Unit 2: Externalities**
- Unit 3: Economic Efficiency and Equity; Social and Private Costs**
- Unit 4: Linear Programming**

Each study unit will take at least two hours, and it include the introduction, objective, main content, self-assessment exercise, conclusion, summary and reference. Other areas border on the Tutor-Marked Assessment (TMA) questions. Some of the self-assessment exercise will necessitate discussion, brainstorming and argument with some of your colleagues. You are advised to do so in order to understand and get acquainted with microeconomics and its application.

There are also textbooks under the references and other (on-line and off-line) resources for further reading. They are meant to give you additional information if only you can lay your hands on any of them. You are required to study the materials; practice the self-assessment exercise and tutor-marked assignment (TMA) questions for greater and in-depth understanding of the course. By doing so, the stated learning objectives of the course would have been achieved.

## Textbook and References

- Adesanjo, W. R., (2014). Impact of Externalities on the health of the Communities in Nigeria, *a Journal Ibadan Printing Press limited*, Vol 4, pp. 22-34.
- Adewole, A. F., (2016). *Introduction to Public Sector Economics*, 2nd edition, Ibadan Press Limited.
- Case, et al. (2014). *Principles of Economics*. (11th ed.). Upper Saddle River, N J: Pearson.
- Chiang, E. (2014). *Core Economics*. (3rd ed.). New York: Worth Publishers.
- Chris B. Murphy (2019), Theory of the Firm). Retrieved on 22/05/2020 from <https://www.investopedia.com/contributors/99607/>
- 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.
- Denardo, E. V. (2002). *The Science of Decision Making: A Problem-Based Approach Using Excel*. New York: John Wiley
- Dixon-Ogbechi, B.N. (2001). *Decision Theory in Business*. Lagos: Philglad Nig. Ltd.
- Dwivedi, D. N. (2002). “*Linear Programming*” *Managerial Economics*. (6th ed.). New Delhi: Vikas Publishing House PVT Ltd.
- Emery, E.D. (1984). *Intermediate Microeconomics*, Harcourt Brace Jovanovich, Publishers, Orlando, USA
- Fashola, M. A. (2009). *Microeconomic theory, Highlights & Policy Extentions for Less-Developed Economies*.
- Francis, A. (1998). “*Linear Equations*” *Business Mathematics and Statistics*. (5th ed.). England: Ashford Colour Press.
- Gary, E. C. (2001). Principles and Practices of Economics. (Interactive student ed.). New York: McGraw-Hill.
- Gupta, P.K. & Hira, D.S. (2012). *Operations Research*. New-Delhi: S. Chand & Company.
- Henderson, J. M., & Quandt, R. E. (2003). *Microeconomic theory: A mathematical approach*. Tata McGraw-Hill
- Jain T.R. & Khana, O.P. (2009). *Managerial Economics*. New Delhi, India:V.K. Publications.
- Jain, T.R. & Khana, O.P. (2009). *Business Economics*. New Delhi, India:V.K. Publications.
- Jhingan M.L (2009): *Microeconomics Theory*, Vrinda Publications (P) Limited, Delhi
- Koutsoyiannis, A. (2001). *Modern Microeconomics* (2nd ed.). Hampshire: Macmillan Press Ltd.
- Krugman, P. & Robin, W. (2012). *Economics*. (3rd ed.). New York: Worth Publishers.
- Lipsey, et al. (2008). Economics. (13th ed.). Upper Saddle River, NJ: Pearson.

Conclusion:

- Lucey, T. (1988). *Quantitative Techniques: An Instructional Manual*. London: DP Publications.
- Luenberger D. G (1973). *Introduction to Linear and Nonlinear Programming*, Addison-Wesley, Reading, Mass.
- Mokhtar S. B., John, J. J. (1977). *Linear Programming and Network Flows*, John Wiley and Sons, New York.
- Nicholson, W. (1978). *Microeconomic Theory: Basic Principles and Extensions*. (2n ed.). Illinois: The Dryden Press.
- Pindyck, R.S & Daniel, L. R. (2009). *Microeconomics*, (7ed), Pearson Education, Inc. New Jersey, USA.
- Robert J. Graham, “How to Compete for Customers: The Bertrand Model of Duopolies in Managerial Economics”. Retrieved on the 22/04/2020 from <https://www.dummies.com/education/economics/how-to-compete-for-customers-the-bertrand-model-of-duopolies-in-managerial-economics/>
- Robert, S. P. & Daniel, L. R. (2009). *Microeconomics*, (7th ed.). New Jersey: Pearson Education International.
- Salvatore, D. (2003). *Microeconomics*, Harper Publishers Inc. New York, USA
- Salvatore, D. (2006). *Microeconomics*. Schuams Series. McGraw-Hill Companies, Inc. USA
- Saqib Shaikh, Distribution and Theories of Distribution (With Diagram) <https://www.economicdiscussion.net/theories-of-distribution/distribution-and-theories-of-distribution-with-diagram/13897/>
- Samia Rekhi, Top 3 Theories of Firm (With Diagram). Retrieved on 22/05/2020 from <https://www.economicdiscussion.net/firm/top-3-theories-of-firm-with-diagram/19519/>
- Shetty, C. M. (1959). Solving Linear Programming Problems with Variable parameters, *Journal of Industrial Engineering*, Vol. 10(6).
- Tucker, I. B. (2014). *Economics for Today*. (8th ed.). Mason, OH: South-Western Cengage Learning.
- Varian, H. R. (n. d.). *Intermediate Microeconomics: A Modern Approach*. (7th ed.).
- Volt, J. W., (2016). Externalities and Manufacturing Industry, *A Journal Macgrill Publication Limited*, Vol 9, pp. 45-57.
- Wash, W. Y. (2017). *Principles and Practice of Public Sector, a textbook*, 1st edition, Macgraw, Publication limited.
- Wright, G. H., (2015). *Principles and Practice of Finance* (1<sup>st</sup> ed.), *Pat and Park Publication limited*.
- Yayah, E. T., (2014). *Introduction to Public Sector Economics*, a textbook 1st edition, Mac and Mix Publication limited.

## Assignment File

Assignment files and marking scheme will be made available to you. This file presents you with details of the work you must submit to your tutor for marking. The marks you



obtain from these assignments shall form part of your final mark for this course. Additional information on assignments will be found in the assignment file and later in this Course Guide in the section on assessment.

There are three assignments in this course. The three course assignments will cover:

Assignment 1 - All TMA questions in Units 1 – 3 (Module 1)

Assignment 2 - All TMA questions in Units 1 – 3 (Module 2)

Assignment 3 - All TMA questions in Units 1 – 4 (Module 3)

### **Assessment**

There are two types of the assessment for the course. First is the tutor-marked assignments; second, is the examination.

In attempting the assignments, you are expected to apply information, knowledge and techniques gathered during the course. The assignments must be submitted to your tutor for formal Assessment in accordance with the deadlines stated in the Presentation Schedule and the Assignments File. The work you submit to your tutor for assessment will count for 30% of your total course mark.

At the end of the course, you will need to sit for a final examination of two hours' duration. This examination will account for 70% of your total course mark.

### **Tutor-Marked Assignments (TMAs)**

There are three tutor-marked assignments in this course. You will submit all the assignments. You are encouraged to work on all the questions thoroughly. The TMAs constitute 30% of the total score.

Assignment questions for the units in this course are contained in the Assignment File. You will be able to complete your assignments from the information and materials contained in your set books, reading 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 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 practice exercises and tutor-marked problems you have previously encountered. All areas of the course will be assessed

Revise the entire course material using the time between finishing the last unit in the module and that of sitting for the final examination. You might find it useful to review

your self-assessment exercises, tutor-marked assignments and comments on them before the examination. The final examination covers information from all parts of the course.

### Course Marking Scheme

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

<b>Assignment</b>	<b>Marks</b>
Assignments (Three assignments submitted and marked)	30%
Final Examination	70%
<b>Total</b>	<b>100%</b>

### Course Overview

The Table presented below indicates the units, number of weeks and assignments to be taken by you to successfully complete the course, Advanced Microeconomic Theory (ECO711).

<b>Units</b>	<b>Title of Work</b>	<b>Week's Activities</b>	<b>Assessment (end of unit)</b>
	Course Guide		
<b>MODULE 1: QUANTITATIVE RELATIONSHIPS, CONSUMER BEHAVIOUR AND EQUILIBRIUM ANALYSIS</b>			
1	<b>Fundamental Quantitative Relationships</b>	Week 1	Assignment 1
2	<b>Theory of Consumer Behaviour</b>	Week 2	Assignment 1
3	<b>Equilibrium Analysis</b>	Week 3	Assignment 1
<b>MODULE 2: PRODUCTION FUNCTIONS, MARKET STRUCTURES, DETERMINATION OF WAGES, RENT, INTEREST AND PROFIT</b>			
1	<b>Production Function, Duopoly and Oligopoly</b>	Week 4	Assignment 2
2	<b>Bilateral Monopoly and Monopsony</b>	Week 5	Assignment 2
3	<b>Theories of Determination of Wages, Rent, Interest and Profit</b>	Week 6	Assignment 2
<b>MODULE 3: THEORIES OF FIRM, DISTRIBUTION AND COST BENEFIT ANALYSIS</b>			
1	<b>Theory of The Firm</b>	Week 7	Assignment 3

2	<b>Theory of Distribution</b>	Week 8	Assignment 3
3	<b>Cost-Benefit Analysis</b>	Week 9	Assignment 3
<b>MODULE 4: WELFARE ECONOMICS AND LINEAR PROGRAMMING</b>			
1	<b>Social Welfare</b>	Week 10	Assignment 4
2	<b>Externalities</b>	Week 11	Assignment 4
3	<b>Economic Efficiency and Equity; Social and Private Costs</b>	Week 12	Assignment 4
4	<b>Linear Programming</b>	Week 13	Assignment 4
	Examination	Week 14 & 15	
	<b>Total</b>	<b>15 Weeks</b>	

### **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 material, and when to embark on discussion with your colleagues. Just as a lecturer might give you an in-class exercise, your study units provides 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 must 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 set books or from a readings 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 twofold. 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-assessments are spread throughout the units, and answers are given at the ends of 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 exercises as you come to it in 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, don't hesitate to call and ask your tutor to provide it.

1. Read this Course Guide thoroughly.
2. Organize 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 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 through each unit.
3. unit.
4. 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 with their course work. If you get into difficulties with your schedule, please let your tutor know before it is too late for help.
5. Turn to Unit 1 and read the introduction and the objectives for the unit.
6. Assemble the study materials. Information about what you need for a unit is given in the 'Overview' 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.
7. 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.
8. Up-to-date course information will be continuously delivered to you at the study centre.
9. 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 exam. Submit all assignments no later than the due date.
10. 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.
11. 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.
12. When you have submitted an assignment to your tutor for marking do not wait for it return before starting on the next units. 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 written on the assignment. Consult your tutor as soon as possible if you have any questions or problems.

13. 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 (1-hours 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 must 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
- You have difficulty with the self-assessment exercises
- You 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**

The course, Advanced Microeconomic Theory (ECO711), exposes you to microeconomics and its applicability and the course guide gives you an overview of what to expect in the course. The course will therefore guide and teach you theories and strategies used by economist analyzing and solving issues upon which public policy decisions are based. The course will therefore be highly important to you as an economist and on the successful completion of the course, you would have developed critical thinking skills with the material necessary for efficient and effective discussion on Advanced Microeconomic Theory.

However, to gain a lot from the course please try to apply anything you learn in the course to term papers writing. We wish you success with the course and hope that you will find it interesting and useful.

# **MODULE 1: QUANTITATIVE RELATIONSHIPS, CONSUMER BEHAVIOUR AND EQUILIBRIUM ANALYSIS**

**Unit 1: Fundamental Quantitative Relationships**

**Unit 2: Theory of Consumer Behaviour**

**Unit 3: Equilibrium Analysis**

## **UNIT 1: FUNDAMENTAL QUANTITATIVE RELATIONSHIPS**

### **CONTENTS**

- 1.0 Introduction
- 2.0 Objectives
- 3.0 Main content
  - 3.1 Quantities and Distances
  - 3.2 Relations between variables: co-ordinate Geometry
  - 3.3 Functional Relations Amongst Variables
  - 3.4 Dependence Amongst Variables
- 4.0 Conclusion
- 5.0 Summary
- 6.0 Tutor Marked Assignment
- 7.0 References /Further Readings

### **1.0 INTRODUCTION**

Economics is concerned with the behaviour of measurable quantities and with the relationship between two or more of these quantities such as wages, prices, national income and employment. This section deals with some of the basic mathematical ideas that are useful in analysing relations between these quantities. Economic laws express the most essential intrinsic relationships and dependencies between economic phenomena, the principal features and trends of economic development. In their interaction with the productive forces, the relations of production are determinate not only qualitatively but also quantitatively.

### **2.0 OBJECTIVES**

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

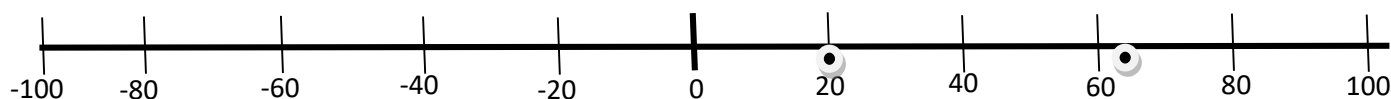
- (i) Understand the relationship between quantities and distances
- (ii) Explain functional relationship and dependency of variables

### **3.0 MAIN CONTENT**

#### **3.1 Quantities and Distances**

If we observe that at a given time, one household buys 20kg of meat while another buys 65kg, we have observed two quantities. It is always possible and very often convenient for us to relate those quantities to distance on a one-dimensional scale. In Fig. 1.1 zero is our point of origin, that is, the point from which all measurement begins. Positive quantities are measured on the right and negative quantities on the left. In Fig 1.1 we have used the following.

**1 cm corresponds to 20kg of meat purchased**



**Figure 1.1: Quantities of meat purchased by the two households**

We can now show the quantities of meat purchased by the two households on the scale in Fig. 1.1 by the dots at 50kg and 90kg.

In the above example we could not have negative quantities because it is impossible for a household to spend less than nothing on meat. However, we could consider a firm that is both selling and buying meat. This firm holds a stock of meat in deep freezers. Sometimes it adds to its stock by buying meat and reduces its stock by selling meat. If we let purchases to be positive numbers (because they add to stocks) and sales to be negative numbers (because they subtract from the stocks), the transaction in which we are interested can range over negative as well as positive quantities. The lower limit is no longer zero but that set by the total stock of meat held since the firm cannot sell more than it has (unless we are dealing in futures market). The firm's purchases and sales can be illustrated in using fig. 1. also.

### **3.2 Relations between variables: co-ordinate Geometry**

In economics we are not only concerned with variable quantities such as a firm's purchases of meat, but also with the relation between two or more measurable quantities, such as meat purchases and the prices of meat. How can we study this relationship? First, we can write the data down in a table as follows:

**Table 1.1**

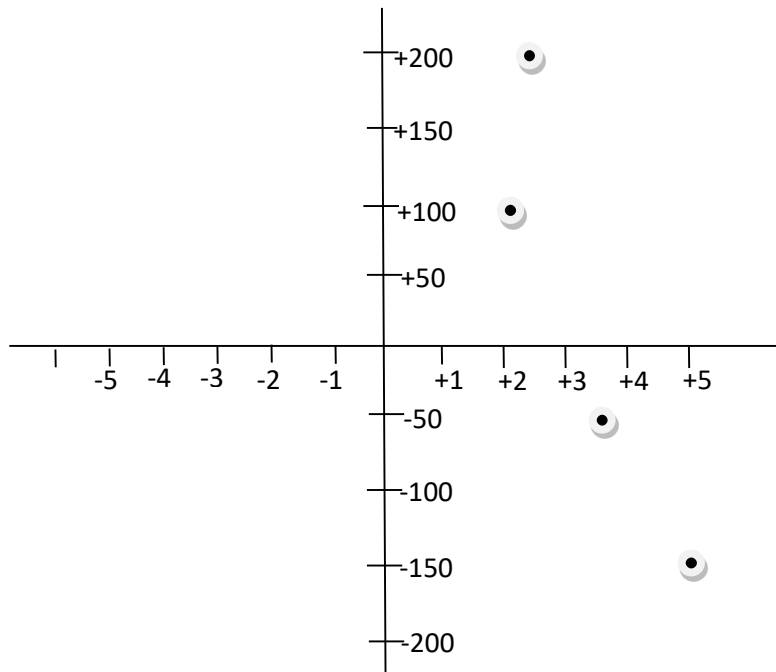
<b>Month</b>	<b>Quantity (kg)</b>	<b>Price</b>
1	- 50	3.7
2	+ 100	2.25
3	+ 200	2.80
4	- 150	5.00

You will notice that in Table 1.1 we have allowed for negative as well as positive quantities. We can also allow for negative prices. (A negative price can be given an economic meaning as a subsidy to the consumer).

Second, we may wish to show the relationship in form of co-ordinate geometry or algebra. Graphically, we can show the price ruling in a particular month and the amount bought or sold in that month by a single dot as in Fig. 1.2.

Setting two lines perpendicular to each other as in the figure divides our total space into four parts, referred to as quadrants. The upper right-hand quadrant is the one in which both variables are positive and is often referred to as the positive quadrant. For every point in the lower right-hand quadrant, y is negative (or zero) and for every point in the lower left-hand quadrant both variables are negative. In many economic problems we can rule out negative values of all the variables as being impossible, and when we draw graphs to show the relationship between such variables, we usually make use of the positive quadrant (as in the case of the usual demand curves)

We are now ready to take the data from Table 1.1 and plot them on a graph.



**Figure 1.2: Quantity of meat bought or sold in a given month**

Each dot in Figure 1.2 shows the quantity of meat bought or sold in a given month, and the price ruling at the time. For instance, dot 4 is 150kg below the origin on the quantity scale and 5.00 to the right of the origin on the price scale, indicating that in the month in which the price was 5.00 the firm sold 150kg of meat.

Thus, we see that the relationship between two variables can be shown on a co-ordinate graph. A single dot indicates simultaneously distances along the two scales which have been set at right angles to each other.

In general, the method of exhibiting a relationship is not confined to that between the price and quantity of meat sold. We can denote any two quantities in which we are interested by  $x$  and  $y$ , and use a graph like Fig. 1.2 to show the relationship between them.

The second point to note, as we have already mentioned, is that the relationship may be explained by either the graphical method or by the use of algebra. Application of a ruler to Fig. 1.2 will reveal that the four points lie on a straight line. What is called “the equation of a straight line” is  $y = ax + b$  ..... (1.1)

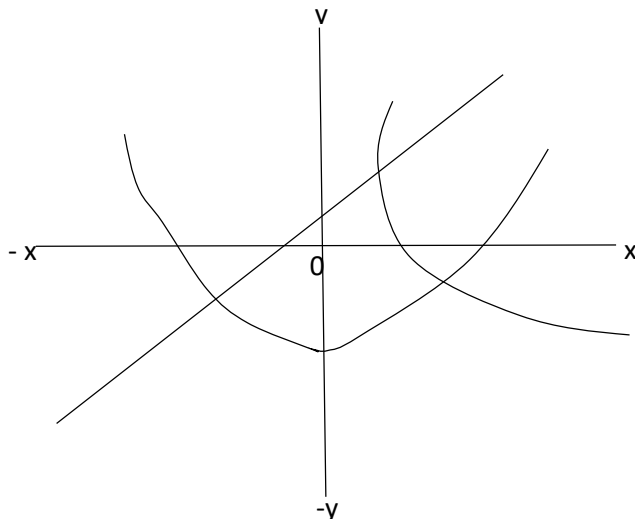
This says that  $y$  is equal to some constant  $b$  plus the value of  $x$  multiplied by a constant  $\alpha$ . Examples of straight lines equations are:

$$y = 2x + 3 \dots\dots\dots (1.2)$$

$$y = 5x - 1 \dots\dots\dots (1.3)$$

The use of co-ordinate geometry is not however confined to straight lines. Let’s look at Fig. 1.3 below. In Fig. 1.3 we have plotted several curves.





**Figure 1.3: Geometric coordinate showing the relationship between x and y**

Each one of them illustrates a possible relationship between x and y, and in each case, if we were given the x-value, we could read-off the corresponding y-value. The equation of some of these curves are quite complex and are not within our course contentment.

### 3.3 Functional Relationships amongst Variables

The idea that there is a relationship between x and y can be expressed quite generally. We may draw illustrative graph as in Figures 1.2 and 1.3; but we may also write:

$$y = f(x)$$

which is stated as ‘y is a function of x’. A graph or function provides us with a rule for getting the value of one variable given the other. When we say that  $y = f(x)$ , we assert that a rule exists, although without saying what it is.

Functions are frequently denoted by Greek letters such as  $\alpha$  and  $\beta$  and other letters from the Roman alphabet such as g. Another notation is to repeat the letter on the left hand side. Thus,  $y = y(x)$  means the same thing as  $y = f(x)$ ; there is a functional relationship between x and y. Economists frequently use this notation because it serves to remind us of the variables we are discussing. Thus, a statement such as ‘investment depends on the rate of interest’ can be written  $I = I(r)$  and ‘consumption depends on the level of income’ as  $C = C(Y)$ .

It is convenient at this point to go through some taxonomy of functions.

(i) **Increasing and Decreasing Functions:** - An increasing function is one the value of which always increases as the value of the independent variable increases. A decreasing function is one the value of which decreases as the independent variable increases. The classes of increasing and decreasing functions are together called monotonic functions, and we may therefore speak of monotonically increasing or monotonically decreasing functions. The straight line.

$$y = ax + b \dots \dots \dots (1.4)$$

provides an example of a monotonic function. If the coefficient  $\alpha$  is positive we have a monotonic increasing function as in

$$y = 2x + 10 \dots \dots \dots (1.5)$$

If the coefficient  $\alpha$  is negative we have a monotonic decreasing function as in

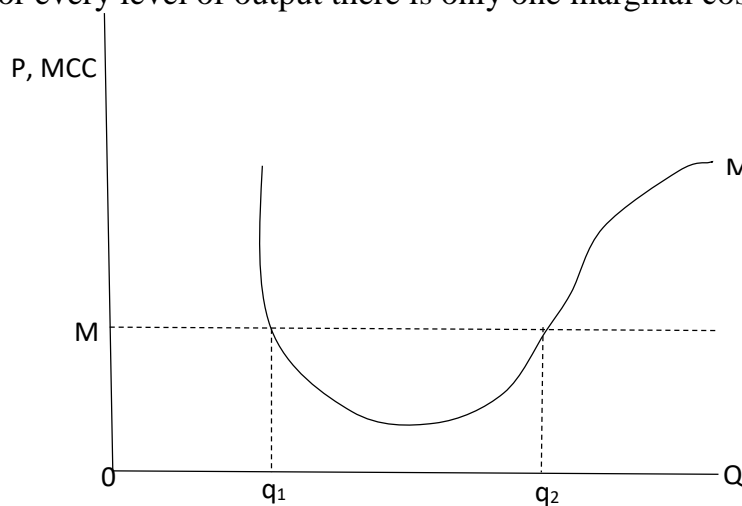
$$y = -2x + 10 \dots \dots \dots (1.6)$$

Many functions increase as  $x$  increases over some range of values and decrease as  $x$  increases over other values. Such functions are not monotonic. A simple example is the quadratic function  $y = ax^2 + bx - c$  ..... (1.7)

Thus, consider  $2x^2 - 3x + 10$  ..... (1.8)

Notice the effect on the value of  $y$  of increasing  $x$  from  $-4$  to  $-3$  and from  $+3$  to  $+4$ . If  $y$  is a single-valued function of  $x$ , then to every value of the variable,  $x$ , there corresponds one and only one value of  $y$ . Similarly, if  $x$  is a single-valued function of  $y$ , then to every value of the variable  $y$  there corresponds one and only one value of  $x$ . The relation is not symmetrical:  $x$  may be a single-valued function of  $y$  without  $y$  being a single-valued function of  $x$ .

Thus in the case a quadratic function in which  $y$  is a single-valued function of  $x$ ,  $x$  is not a single-valued function of  $y$ : two  $x$ -values map into each  $y$ -value. This symmetry occurs frequently in the functions we deal with in economics. Thus, consider a U-shaped marginal cost curve. Each level of output gives a unique level of marginal cost, but to a value of marginal cost there corresponds two levels of output. For instance, in Figure 1.4, the MC, Corresponds to two levels of output at  $q_1$  and  $q_2$ . It is also clear from inspection that for every level of output there is only one marginal cost.



**Figure 1.4: Marginal Cost Curve**

### 3.4 Dependence amongst Variables

The existence of a function or mapping rule, from one variable to another does not however imply anything about causation. If we wish to deal analytically with a causal relationship, we shall obviously write it in a functional form, but the assertion that there exists a mapping rule between two variables says nothing about causes.

When we write  $y = f(x)$ , we are only saying, let  $y$  be the value obtained when we operate on  $x$  by the rule  $f$ . If we also suppose that  $x$  stands for some observable quantity and that  $y$  stands for another observable quantity, then when we write  $y = f(x)$  we only say 'the value of  $y$  can be found by operating on  $x$  in the manner directed'. This means that  $x$  and  $y$  are systematically associated. This systematic association may be because  $x$  influences  $y$ , because  $y$  influences  $x$ , because although  $x$  and  $y$  are not related they are both influenced by some common causal factor  $z$ , or because  $y$  and  $x$  just happen to have varied with, each other by chance.

Consider the following functional relationships.

$$y = y(x) \dots \dots \dots (1.9)$$

$$x = x(y) \dots \dots \dots (1.10)$$

The term on the left-hand side of each of equations (1.9) and (1.10) is called the dependent variable, and the terms on the right-hand side are called independent variables (or variable, if there is only one). As far as mathematics is concerned, the distinction dependent and independent variable is arbitrary:  $y = y(x)$  implies  $x = x(y)$ . The convention may be used, however, to express information we have about the causal relation between the variables. Assume for example, that crop yield  $C$  depends solely on the amount of rainfall,  $R$ . This allows us to write.

$$C = C(R) \dots \dots \dots (1.11)$$

If, however, we can deduce the amount of crop if we know the rainfall then we must be able to deduce the amount of rainfall if we know the crop. Thus we also have

$$R = R(C) \dots \dots \dots (1.12)$$

As far as mathematics is concerned, it does not matter which of these two forms we choose, (1.11) or (1.12). But, of course, the causal relation is clearly defined in this case. The amount of rainfall influences the crop yield; the crop yield does not influence the amount of rainfall.

As a matter of convention, wherever we think we know the direction of the causal link between two variables we write the causes as independent variables and the effects as dependent ones. Thus, it would be consistent without ideas about the physical facts to write equation (III) instead of equation (1.12). Again if we wanted to say that crop yield  $C$  depended on fertilizer,  $F$ , sunshine,  $S$ , and rainfall,  $R$ , we would write.

$$C = C(F, S, R) \dots \dots \dots (1.13)$$

In equation (1.13)  $C$  is the dependent variable and  $F$ ,  $S$ , and  $R$  are the independent variable. Another example of a function of more than one variable in economics is

$$P = MV/T \dots \dots \dots (1.14)$$

Where  $P$  is the price level  $M$  is the quantity of money,  $V$  is the velocity of circulation of money, and  $T$  is the total volume of transactions. Equation (1.14) specifies the exact relationship between the dependent variable,  $P$ , and each of the independent variables,  $M$ ,  $V$  and  $T$ . If we wished to say merely that the price level was a function of these three variables without committing ourselves to the exact form of the relation in the way that we did above, we could merely write:

$$P = P(M, V, T) \dots \dots \dots (1.15)$$

This is a convenient moment to introduce another terminology. The variables that appear in brackets after the functional symbol are often referred to as the arguments of the function. Arguments and independent variables are nearly, but not quite, the same thing, and the distinction can be important. Suppose that we have  $y = f(x/z)$  and compare this with  $y = g(x, z)$ . In both functions,  $x$  and  $z$  appear as the independent variables. In the first case, however, their ratio appears explicitly, whereas in the second case they appear separated by a comma. What this means is that  $f$  is an unspecified mapping rule from the ratio  $x/z$ , whereas  $g$  is some unspecified mapping rule from pairs

of values of  $x$  and  $z$  entered separately. In the case of  $f(x/z)$  the value of neither variable matters for itself, but only as it affects  $x/z$ . Hence in this case we say that  $x/z$  is the argument of the function, while in the case of  $g(x, z)$  we say that both  $x$  and  $z$  are arguments. Obviously, if we are dealing with causal relationships, it is a matter of some importance to be clear about what the arguments of the functions are. This is because functions are mapping rules from their arguments to the dependent variable.

In drawing graphs of functions, it is the usual convention to plot the independent variable on the  $x$ -axis and the dependent variable on the  $y$ -axis. However, Alfred Marshall decided to reverse this practice in the case of demand and supply curves. He put the independent variable, price, on the  $y$ -axis and the dependent variable, quantity, on the  $x$ -axis. This practice has persisted to this day to the confusion of students, although in the case of every other diagram the normal practice is adopted with the dependent variable on the  $y$ -axis and the independent variable on the  $x$ -axis. Thus, although there is no hard and fast rule about it, you should be prepared, as a matter of general practice, to redraw any familiar economic diagram with the axes reversed.

#### **4.0 CONCLUSION**

Fundamental quantitative relationships provide analysts with the tools to examine and analyse past, current, and anticipated future events. Any subject involving numbers can be quantified; there are thus many fields in which quantitative analyses are used. This has some benefits.

#### **5.0 SUMMARY**

Fundamental quantitative relationships are the representation and analysis of information about the direction of change (+, -, or 0) in some economic variable(s) as related to change of some other economic variable(s). For the non-zero case, what makes the change quantitative is its direction and not its magnitude which is specified.

#### **6.0 TUTOR-MARKED ASSIGNMENT**

1. Explain what is meant by the term, *quantitative relationships* and how it relates to distances, functional relationship and dependency of variable.
2. Identify the variable(s) involved in the question below stating which is the dependant variable and other wise.
  - a. What is the relationship between gender and O'Level exam results amongst NOUN students?
  - b. What is the relationship between disposable income and location amongst graduates?
  - c. What is the relationship between job satisfaction and salary amongst Nigerian residents?

#### **7.0 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. (2001). *Modern Microeconomics*, (2nd ed.). Hampshire: Macmillan Press Ltd.
- Robert, S. P. & Daniel, L. R. (2009). *Microeconomics*, (7th ed.). New Jersey: Pearson Education International.

## **UNIT 2: THEORY OF CONSUMER BEHAVIOUR**

### **CONTENTS**

- 1.0 Introduction
- 2.0 Objectives
- 3.0 Main Content
  - 3.1 Utility Analysis
  - 3.2 The Cardinal Utility Approach
    - 3.2.1 Definition of Concepts
    - 3.2.2 Relationship between Total Utility and Marginal Utility
    - 3.2.3 Equilibrium of the consumer in the Cardinal Utility Approach
    - 3.2.4 Derivation of Equilibrium
    - 3.2.5 Assumptions of Cardinal Utility Approach
  - 3.3 The Ordinal Utility Approach
    - 3.3.1 Assumption of the Ordinal Utility Approach
    - 3.3.2 Indifference Curve
    - 3.3.3 Properties of Indifference Curve
    - 3.3.4 Marginal Rate of Substitution
    - 3.3.5 Budget line
    - 3.3.6 Consumer Equilibrium
  - 3.4 Revealed Preference Approach
    - 3.4.1 Assumptions of Revealed Preference
- 4.0 Conclusion
- 5.0 Summary
- 6.0 Tutor Marked Assignment
- 7.0 References /Further Readings

### **1.0 Introduction**

A basic assumption of microeconomics is that because a consumer has a limited budget, spending must be judiciously allocated for maximum benefit. Microeconomics also supposes that consumers make their buying decisions in an effort to get the most satisfaction at the least cost - in other words, maximising utility.

It is worthy to note that the consumer is considered to be a rational person, who spends in order to derive the maximum amount of satisfaction, or utility, from it. An individual consumer wants to get all to maximise their total utility.

Consumer behaviour is also the study of consumers, groups, or organisations and the processes they use to select, secure, and dispose of goods, services, experiences, or ideas to satisfy needs and the effects that these processes have on the consumer and society. It is how consumers allocate their money incomes among goods and services. The basis of consumer behaviour is underlined by the thinking referred to as law of diminishing marginal utility (LDMU).

### **2.0 Objectives**

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

- understand the utility concepts
- know the relationship between total utility and marginal utility
- differentiate between cardinal and ordinal utility theory

- explain the assumptions of Revealed Preference Hypothesis and indifference curves
- explain the consumer's equilibrium and its conditions in both the cardinal and ordinal utility theory
- analyse the Revealed Preference Hypothesis and indifference curves

### 3.0 Main Content

#### 3.1 Utility Analysis

Utility analysis simply investigates consumer behaviour and is focused on the satisfaction of wants and needs obtained by the consumption of goods by a consumer. This is technically termed utility. There are several ways of analyzing the utility a consumer derives from consuming any commodity. In the context of this course, we shall confine our analysis to the following;

- ❖ The Cardinal Utility Approach
- ❖ The Ordinal Utility Approach
- ❖ The Revealed Preference Approach

#### 3.2 The Cardinal Utility Approach

The cardinal utility approach believe that utility can be quantified. i.e. the amount of satisfaction a consumer derives from consuming a particular commodity can be measured. The scale of measurement is “utils”

##### 3.2.1 Definition of Concepts

**a. Total Utility (TU):** This is the summation of all the satisfaction derived from consuming a particular commodity. TU depends on quantity consumed. i.e.  $TU = f(q_1, q_2, q_3, \dots, q_n)$

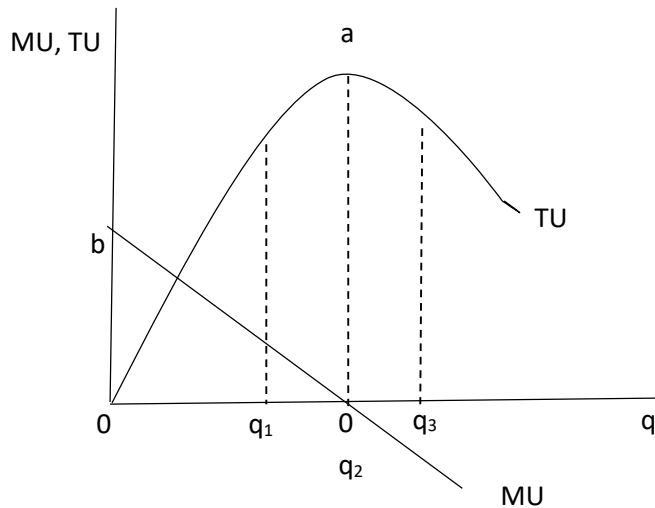
**b. Average Utility (AU):** This is defined as total utility derived by quantity consumed. i.e.  $AU = \frac{TU}{q}$

**c. Marginal Utility (MU):** This is the additional satisfaction a consumer derives from consuming an extra unit of a particular commodity. It can also be defined as the slope of the utility function with respect to quantity consumed. i.e.  $MU = \frac{dTU}{dq}$  or  $MU = \frac{dU}{dq}$

**d. Diminishing Marginal Utility (MU):** The law of DMU states that “as a consumer acquires successive units of a given commodity, the additional satisfaction falls”.

##### 3.2.2 Relationship between Total Utility and Marginal Utility

As quantity consumed of a particular good is increasing, total utility (TU) will be increasing, marginal utility (MU) will be falling indicating that the increase in TU is at a declining rate. At the maximum point of total utility, marginal utility is zero. Quantities beyond this point where total utility is maximized will lead to a negative marginal utility (MU).



**Figure 2.1: Relationship between Total Utility and Marginal Utility**

The figure above shows the relationship between total utility (TU) and marginal utility (MU). As the consumer begins to consume good  $q$ , TU started to rise. As the consumer consumed  $q_1$  quantity of goods, TU is still rising, while MU is falling, showing that TU is rising at a declining rate. At  $q_2$ , TU is maximized while MU is zero (0). At  $q_3$ , TU falls while MU is less than zero, i.e. negative

**Example 1**

Given a utility function  $U = 144q - 3q^2$ , find the

- i. Marginal Utility (MU)
- ii. Utility maximizing point

**Solution**

i.  $MU = \frac{dU}{dq} = 144 - 6q$

The Marginal Utility =  $144 - 6q$

- ii. The utility maximizing point is the point where  $MU = 0$ , so we equate  $MU = 0$

$$\begin{aligned} \text{Therefore, } 144 - 6q &= 0 \\ &= 144 = 6q \\ q &= \frac{144}{6} \\ q &= 24 \end{aligned}$$

**Example 2**

Assuming the utility function  $U = 18q + 7q^2 - \frac{1}{3}q^3$ , determine

- i. The utility maximizing quantity
- ii. At what unit of consumption does the law of diminishing marginal utility (DMU) set in?

**Solution**

i.  $MU = \frac{dU}{dq} = 18q + 14q - q^2$

At this point, we apply the quadratic formula

$$q = \frac{-(-14) \pm \sqrt{14^2 - 4 \times 1 \times -18}}{2(1)}$$

$$q = \frac{14 \pm 16.37}{2}$$

Since we are dealing with quantity, we go with the positive value  
 $q = 15.9$

ii. Diminishing marginal utility sets in as soon as  $MU = 0$

Remember that  $MU = 18q + 14q - q^2$

Therefore  $18q + 14q - q^2 = 0$

To get the unit at which the law of DMU set in, we maximize MU and set to zero

$$\begin{aligned} \frac{dMU}{dq} = 0 &= 14 - 2q \\ 14 &= 2q \\ q &= 7 \end{aligned}$$

Thus, at 7 units, DMU sets in

### 3.2.3 Equilibrium of the Consumer in the Cardinal Utility Approach

A consumer is at equilibrium if and only if utility is maximized. The condition for utility maximization in a single commodity case is that  $MU = Price$ , i.e. the additional satisfaction a consumer derive from consuming a particular commodity must be equal to its price.

$$\begin{aligned} MU = P &= \text{Equilibrium} \\ MU < P &= \text{Disequilibrium} \\ MU > P &= \text{Disequilibrium} \end{aligned}$$

If  $MU < P$ , to restore equilibrium, a consumer should consume more of the commodity and should continue to do so until  $MU = P$ .

If  $MU > P$ , to restore equilibrium, a consumer should cut down the consumption of the commodity and should continue to do so until  $MU = P$ .

For multi-commodity case equilibrium is attained when the ratios of MU to price are equal for all commodities. i.e.

$$\frac{MU_1}{P_1} = \frac{MU_2}{P_2} = \frac{MU_3}{P_3} = \dots = \frac{MU_n}{P_n} \text{ --- Equi - Marginal Utility}$$

### 3.2.4 Derivation of Equilibrium

**For a single commodity case**

$$\text{Max } U = f(q) \text{ --- (1)}$$

Subject to  $Y = pq$  --- (2), note that equation 2 is the consumer's income

The consumer will maximize the difference between his utility and expenditure function.



Thus we subtract equation (2) from equation (1)

$$L = U - pq \text{ --- (3)}$$

$$\frac{dL}{dq} = \frac{dU}{dq} - p = 0 \text{ --- (4)}$$

Working with the right hand side of equation (4)

$$\frac{dU}{dq} - p = 0$$

But recall that  $\frac{dU}{dq} = MU$

Therefore equation (4) becomes

$$MU - p = 0$$

$$MU = p$$

Therefore, at equilibrium, *Marginal Utility (MU) = Price(p)*

### Derivation of Equilibrium in Multi commodity case (Equi Marginal Utility)

Given that  $U = f(q_1, q_2, \dots, q_n) \text{ --- (1)}$

Subject to  $Y = p_1q_1 + p_2q_2 + \dots + p_nq_n \text{ --- (2)}$ , Note that equation (2) is the constraint equation

Rewriting the constraint equation and setting to zero, we have

$$Y - p_1q_1 - p_2q_2 - \dots - p_nq_n = 0 \text{ --- (3)}$$

Introducing the langrage multiplier; a strategy used to find a local maxima and minima of a function subject to a constraint, we thus multiply equation (3) by  $\lambda$

$$\lambda(Y - p_1q_1 - p_2q_2 - \dots - p_nq_n) = 0 \text{ --- (4)}$$

Equation (1) + Equation (4)

$$Z = U + \lambda(Y - p_1q_1 - p_2q_2 - \dots - p_nq_n) = 0 \text{ --- (5)}$$

Next we maximize equation (5) and set to zero. i.e.

$$\frac{dZ}{dq_1} = 0, \quad \frac{dZ}{dq_2} = 0, \quad \frac{dZ}{dq_n} = 0$$

$$\frac{dZ}{dq_1} = \frac{dU}{dq_1} - \lambda p_1 = 0 \text{ --- (6)}$$

$$\frac{dZ}{dq_2} = \frac{dU}{dq_2} - \lambda p_2 = 0 \text{ --- (7)}$$

$$\frac{dZ}{dq_n} = \frac{dU}{dq_n} - \lambda p_n = 0 \text{ --- (8)}$$

From equations (6), (7) and (8) we have

$$\frac{dU}{dq_1} = \lambda p_1, \frac{dU}{dq_2} = \lambda p_2, \frac{dU}{dq_n} = \lambda p_n \text{ --- (9)}$$

Recall that,  $\frac{dU}{dq_1} = MU_1$

Equation 9 becomes

$$\lambda = \frac{MU_1}{p_1} = \frac{MU_2}{p_2} = \frac{MU_n}{p_n} \text{ --- --- --- Equi Marginal Utility}$$

### 3.2.5 Assumptions of Cardinal Utility Approach

1. Utility is measurable and quantifiable: i.e. it is possible for the consumer to quantify the amount of utils being derived from consuming a commodity. This assumption enables one to compare utility with respect to size and the level of satisfaction derivable from consuming a commodity with that of another
2. Utility is additive: This means that the amount of utility derived from a commodity is a function of that commodity alone and independent of the quantity consumed of any other good. Hence the total utility a consumer derives from a given basket of goods and services is the sum of total of the separate utilities derived from goods and services content in the basket alone.
3. The marginal utility of money is constant: While the marginal utility of goods and services diminishes as more and more units of the goods is consumed, the marginal utility of money spent on the goods remains constant although the amount of money left with the consumer declines
4. Total utility depends on quantity consumed
5. Diminishing marginal utility

### SELF ASSESSMENT EXERCISE

With illustration, discuss the equi marginal principle

### 3.3 The Ordinal Utility Approach

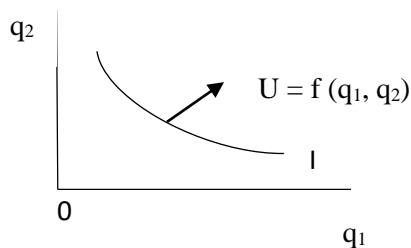
The ordinal approach believe that utility cannot be measured, but that an individual can rank the various bundles available to him in their order of preference. It is the use of indifference curves to evaluate preferences in order of most preferred to least preferred but does not concern itself with how much one combination is preferred to another.

#### 3.3.1 Assumption of the Ordinal Utility Approach

- ✓ Rationality
- ✓ Ordinality
- ✓ Consistency
- ✓ Transitivity
- ✓ Divisibility
- ✓ Reflexivity
- ✓ Diminishing Marginal Rate of Substitution
- ✓ Continuity
- ✓ Non-Satiation
- ✓ Convexity

### 3.3.2 Indifference Curve

An indifference curve is the locus of all commodity combinations that yield the same level of satisfaction. By assuming that 2 commodities are consumed ( $q_1$  and  $q_2$ ), the indifference curve can be defined as the locus of combination of  $q_1$  and  $q_2$  that yields the same level of satisfaction.



**Figure 2.2: An Indifference Curve**

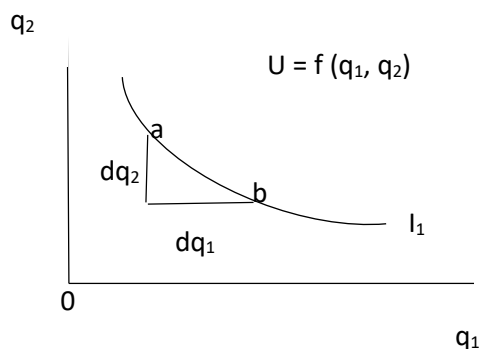
The diagram above shows an indifference curve ( $I_1$ ) indicating various points of combination of  $q_1$  and  $q_2$  that yield the same level of satisfaction

### 3.3.3 Properties of Indifference Curve

- ✓ Negatively inclined. i.e. downward sloping
- ✓ Convex to the origin
- ✓ Direction of increase in utility is from north-east
- ✓ Densed indifference map
- ✓ Indifference curves do not cut or intersect one another
- ✓ Indifference curves do not touch the axis

### 3.3.4 Marginal Rate of Substitution (MRS)

This is the slope of the indifference curve. It is the rate at which one commodity can be substituted for the other. It can also be defined as the ratio of the marginal utility of the commodities consumed.



**Figure 2.3: Marginal Rate of Substitution (MRS)**

$$\frac{-dq_2}{dq_1} = MRS_{1,2} = \frac{MU_1}{MU_2}$$

### 3.3.5 Budget line

This is defined as the locus of all commodity combinations that can be purchased with the same amount of money income. In terms  $q_1$  and  $q_2$ , the budget line is the locus of points of combinations of  $q_1$  and  $q_2$  that can be repurchased with the same amount of money income  $Y$  (fixed). Thus the value of  $Y$  is given as  $Y = p_1q_1 + p_2q_2$

$$\text{Thus } p_1q_1 = Y - p_2q_2$$

$$q_1 = \frac{Y - p_2q_2}{p_1}$$

Also

$$p_2q_2 = Y - p_1q_1$$

$$q_2 = \frac{Y - p_1q_1}{p_2}$$

Assuming the consumer buys only  $q_1$ , thus  $q_2 = 0$ . The maximum  $q_1$  bought will be

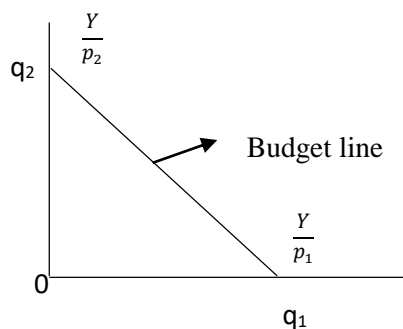
$$q_1 = \frac{Y - p_2 \cdot 0}{p_1}$$

$$q_1 = \frac{Y}{p_1}$$

Similarly, if the consumer buys only  $q_2$ , thus  $q_1 = 0$ . The maximum  $q_2$  bought will be

$$q_2 = \frac{Y - p_1 \cdot 0}{p_2}$$

$$q_2 = \frac{Y}{p_2}$$



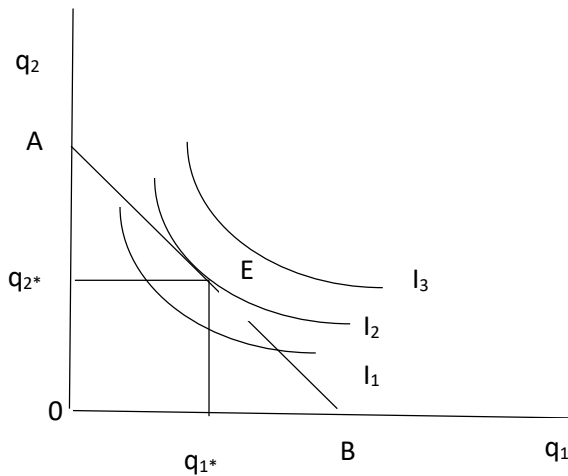
**Figure 2.4: Budget Line**

### 3.3.6 Consumer Equilibrium

The consumer is at equilibrium if and only if utility is maximized. The conditions for utility maximization can be stated as follows

First Order Condition: The slope of the indifference curve must be equal to the slope of the budget lines. i.e  $MRS_{1,2} = \frac{p_1}{p_2}$ .

Second Order Condition: The indifference curve must be convex to the origin. This is mathematically stated as the Bordered Hessian Determinant (BHD) must be positive.



**Figure 2.5: Graphical Illustration of Consumer Equilibrium**

In the figure above, AB is the budget line,  $I_1$ ,  $I_2$  and  $I_3$  are indifference curve representing different levels of satisfaction. AB can afford  $I_1$  but the consumer being rational will not operate along  $I_1$  because AB can afford a higher level of utility.  $I_3$  is unattainable or not feasible given the budget line AB. The rational and feasible level of utility is  $I_2$ .  $I_2$  is tangential to AB at point E. given their tangency at point E, their slope of  $AB = \frac{p_1}{p_2}$ . Thus, at E,  $MRS_{1,2} = \frac{p_1}{p_2}$  satisfying the first order condition. The second order condition is also satisfied because the indifference curve is convex to the origin.

First order condition is necessary but not sufficient for equilibrium because if we have an indifference curve that is tangential to the budget line but not convex to the origin, first order condition is satisfied but second order condition is not satisfied.

Second order condition is sufficient but not necessary for equilibrium because if what have an indifference curve that is convex to the origin but not tangential to any budget line, second order condition is satisfied but first order condition is not satisfied.

Thus for equilibrium to be attained, the necessary and sufficient conditions are required.

### 3.4 Revealed Preference Approach

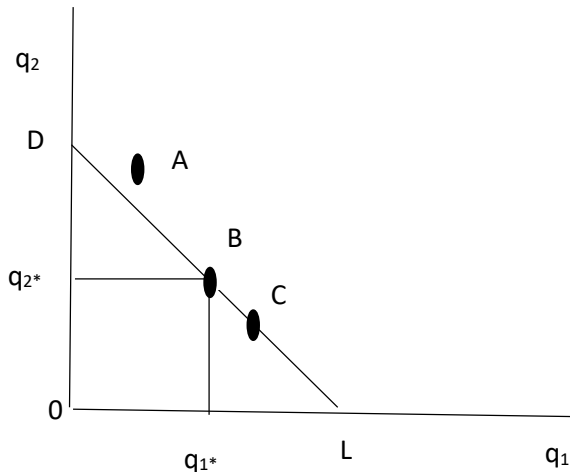
This approach developed by Paul Samuelson is based on the notion that choice reveals preference. i.e. if a consumer chooses a particular alternative in the face of other alternatives given a budget situation, the chosen alternative is revealed to be preferred to other alternatives.

If there is A, B and C as bundles of commodities, if A is chosen, it means that A is preferred to B and A is preferred to C.

### 3.4.1 Assumptions of Revealed Preference

- ✓ Rationality
- ✓ Choice reveals preference
- ✓ Consistency
- ✓ Transitivity

### Theory of Revealed Preference



**Figure 2.6: Revealed Preference Theory**

From the figure above, assuming the consumer chooses point B, it means that B is preferred to point A and C, i.e. point A and C are inferior to point B. Thus B is revealed to be preferred to A and C and every other point along DL.

### 4.0 CONCLUSION

- According to the cardinal theory, utility can be quantified in terms of the money a consumer is willing to pay for it.
- Ordinal theory stated that utility cannot be measured in terms of monetary value due to changes in the utility of money.
- In the cardinal theory, a consumer is at equilibrium if and only if utility is maximized. Utility is maximized at that point where  $MU_x = P_x$
- In the ordinal theory, a consumer is at equilibrium if and only if utility is maximized based on certain conditions. The conditions include that the slope of the indifference curve must be equal to the slope of the budget lines. i.e.  $MRS_{1,2} = \frac{p_1}{p_2}$ , and the indifference curve must be convex to the origin.
- Revealed preference hypothesis stated 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.
- Higher indifference curve gives a higher satisfaction.

## 5.0 SUMMARY

In this unit, you learnt about the theory of consumer behaviour and basic assumptions underlining the theory of consumer behaviour. It also introduced you to the law of diminishing marginal utility (LDMU), the approaches in analyzing consumer behavior, the equilibrium point of a consumer and necessary conditions to attain equilibrium.

## 6.0 TUTOR-MARKED ASSIGNMENT

1. Given a utility function  $U = 81q - 4q^2$ , find the
  - i. Marginal Utility (MU)
  - ii. Utility maximizing point
2. Derive the equilibrium in Multi commodity case in the cardinal utility theory
3. Using graphical illustration, discuss the consumer equilibrium in the ordinal utility theory

## 7.0 REFERENCES/FURTHER READING

Henderson, J. M., & Quandt, R. E. (2003). *Microeconomic theory: A mathematical approach*. Tata McGraw-Hill

Jhingan M.L (2009): *Microeconomics Theory*, Vrinda Publications (P) Limited, Delhi

Koutsoyiannis, A. (2001). *Modern Microeconomics* (2nd ed.). Hampshire: Macmillan Press Ltd.

# UNIT 3: EQUILIBRIUM ANALYSIS: PARTIAL EQUILIBRIUM, GENERAL EQUILIBRIUM AND DYNAMIC EQUILIBRIUM ANALYSIS

## CONTENTS

- 1.0 Introduction
- 2.0 Objectives
- 3.0 Main content
  - 3.1 Partial Equilibrium
  - 3.2 General Equilibrium
    - 3.2.1 Assumptions
    - 3.2.2 Static Properties of a General Equilibrium State
  - 3.3 The Walrasian System: Algebraic Approach
    - 3.3.1 Equations of the Model
  - 3.4 Dynamic Equilibrium
    - 3.4.1 Micro-Dynamics (cobweb)
      - 3.3.1.2 The Cob-web Model
    - 3.3.2 Macro-Dynamics (cobweb)
- 4.0 Conclusion
- 5.0 Summary
- 6.0 Tutor Marked Assignment
- 7.0 References /Further Readings

## 1.0 INTRODUCTION

In this module and the units that follow, our objective is to use the tools of microeconomic analysis to develop a system-wide view of the allocation of resources and output in a market economy. The principal question to be approached is whether, in a competitive market economy, the individual actions of microeconomic agents (consumers, firms) can result in a general, system-wide equilibrium such that all product and factor prices are consistent with equality between quantity demanded and quantity supplied in each separate market.

## 2.0 OBJECTIVES

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

1. Differentiate between the concept of general equilibrium and partial equilibrium
2. Explain dynamic equilibrium.

## 3.0 MAIN CONTENT

### 3.1 Partial Equilibrium

Partial equilibrium as the name suggests, takes into consideration, only a part of the market to attain equilibrium, *ceteris paribus*.

A partial equilibrium is based on only a restricted range of data. Example include, the price of a single product, the prices of all other products being held fixed during the analysis. Partial equilibrium analysis examines the effects of policy action in creating equilibrium only in that particular sector or market which is directly affected, ignoring its effect on any other market or industry assuming that their impact is small (if any). A partial equilibrium model is an economic model of a single market. It is the study of the



behaviour of individual decision-making units and of the workings of individual markets, viewed in isolation.

Partial equilibrium analysis is the analysis of an equilibrium position for a sector of the economy or for one or several partial groups of the economic unit corresponding to a particular set of data. Partial or particular equilibrium analysis, also known as microeconomic analysis, is the study of the equilibrium position of an individual, a firm, an industry or a group of industries viewed in isolation. In other words, this method considers the changes in one or two variables keeping all others constant, *ceteris paribus* (others remaining the same). The *ceteris paribus* assumption is the crux of partial equilibrium analysis. Let's look at the following examples:

- (a) **Consumer's Equilibrium:** With the application of partial equilibrium analysis, a consumer's equilibrium is indicated when he is getting maximum satisfaction from a given expenditure and in a given set of conditions relating to price and supply of the commodity.

The conditions are:

- (1) the marginal utility of each good is equal to its price (P):

$$\frac{MU_A}{P_A} = \frac{MU_B}{P_B} = \dots = \frac{MU_N}{P_N};$$

- (2) the consumer must spend his entire income (Y) on the purchase of goods:

$$Y = P_A Q_A + P_B Q_B + \dots + P_N Q_N.$$

It is assumed that his tastes, preferences, money income and the prices of the goods are constant.

- (b) **Producer's Equilibrium:** A producer is in equilibrium when he is able to maximise his aggregate net profit in the economic conditions in which he is working.
- (c) **Firm's Equilibrium:** A firm is said to be in long-run equilibrium when it has attained the optimum size which is ideal from the viewpoint of profit and utilization of resources at its disposal.
- (d) **Industry's Equilibrium:** Equilibrium of an industry shows that there is no incentive for new firms to enter it or for the existing firms to leave it. This will happen when the marginal firm in the industry is making only normal profits. In all these cases, those who have incentive to change it have no opportunity and those who have the opportunity have no incentive.

### Assumptions

1. Commodity price is given for the consumers.
2. Consumer's taste and preferences, habits, incomes are also considered to be constant.
3. Prices of prolific resources of a commodity and that of other related goods (substitute or complimentary) are known as well as constant.
4. Industry is easily availed with factors of production at a known and constant price compliant with the methods of production in use.

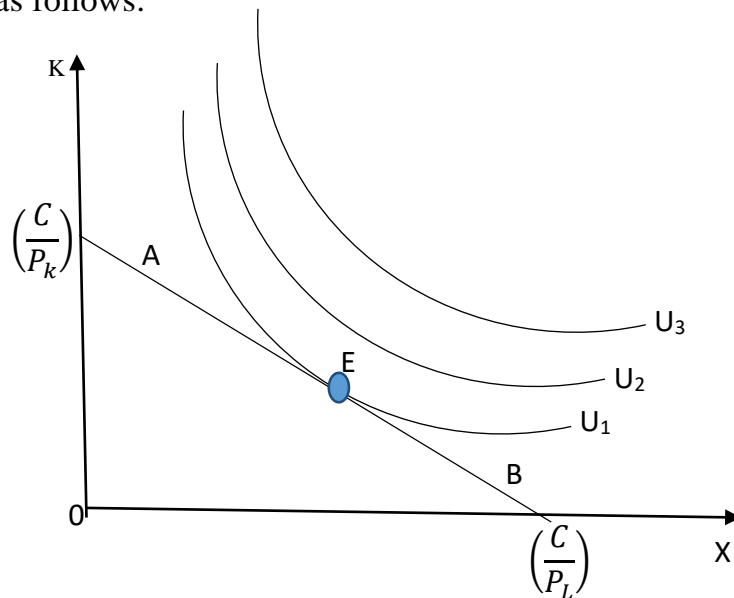
5. Prices of the products that the factor of production helps in producing and the price and quantity of other factors are known and constant.
6. There is perfect mobility of factors of production between occupation and places.

### 3.1.1 Partial Equilibrium Conditions

Let us first review partial equilibrium conditions in consumption and production.

(a) **Consumption:** Recall we stated in the theory of consumers behaviour that equilibrium or optimum consumption is said to be attained, in a world of two goods, x and y, given the society's indifference map, the prices of x and y, and the income of the consumer, when the slope of the indifference curve  $\left(-\frac{dy}{dx} = MRS_{x'y}\right)$  is equal to the ratio of the prices of commodity  $\left(\frac{P_x}{P_y}\right)$ .

(b) **Production:** given the production isoquant map, two factors of production, labour and capital, and factor prices,  $P_L$  and  $P_K$ , equilibrium in production is attained when the slopes of the isoquant curve and the isocost curve are equal. This is shown graphically as follows:



**Figure 3.1: Production Equilibrium**

At E the slope of the isocost line  $AB \frac{\frac{C}{P_k}}{\frac{C}{P_L}} = \frac{C}{P_k} \times \frac{P_L}{C} = \frac{P_L}{P_k}$  is equal to the slope of the isoquant curve. A movement downwards along the isoquant curve shows the increasing difficulty of substituting K (capital) for L (Labour) and is written as  $\frac{dk}{dl} = \frac{\partial x \partial x \partial k}{\partial l} =$

$$\frac{M^P_L}{M^P_K} \frac{dl}{dk} = \frac{\frac{dx}{dl}}{\frac{dx}{dk}} = \frac{MP_L}{MP_K} = MRTS_{L,K}$$

Thus equilibrium holds when  $\frac{P_L}{P_K} = MRTS_{L,K} \dots \dots \dots (3.2)$

where  $MRTS_{L,K}$  is the marginal rate of technical substitution between labour and capital. This condition applies to the production of x. For equilibrium to hold in the case of y, the same condition must be satisfied.

## SELF-ASSESSMENT EXERCISE

What is partial equilibrium analysis?

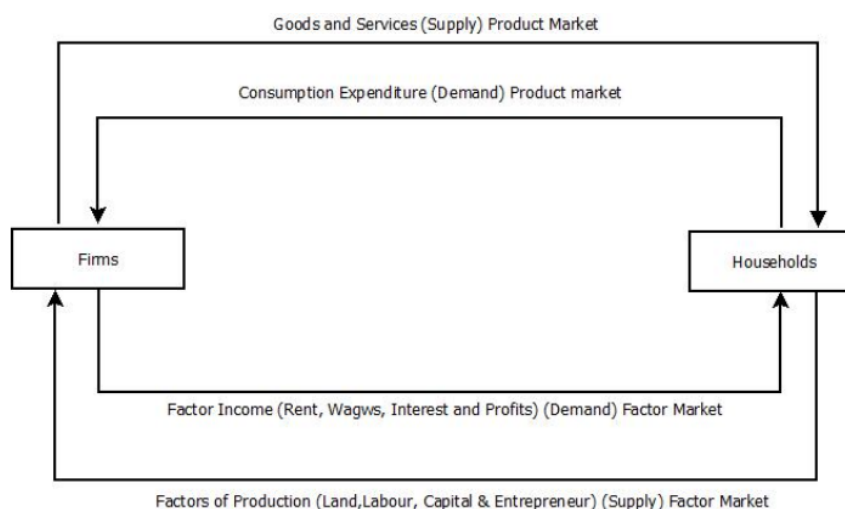
### 3.2 General Equilibrium

General equilibrium analysis studies the behaviour of all individual decision-making units and of all individual markets simultaneously. At this point, we turn to general equilibrium analysis to describe the operation of an economic system made up of many markets.

The partial equilibrium analysis studies the relationship between only selected few variables, keeping others unchanged. Whereas the general equilibrium analysis enables us to study the behaviour of economic variables taking full account of the interaction between those variables and the rest of the economy. In partial equilibrium analysis, the determination of the price of a good is simplified by just looking at the price of one good and assuming that the prices of all other goods remain constant. Thus the economy is in general equilibrium when commodity prices make each demand equal to its supply and factor prices make the demand for each factor equal to its supply so that all product markets and factor markets are simultaneously in equilibrium. Such a general equilibrium is characterized by two conditions in which the set of prices in all product and factor markets is such that:

- 1) All consumers maximize their satisfactions and all producers maximize their profits and
- 2) All markets are cleared which means that the total amount demanded equals the total amount supplied at a positive price in both the product and factor markets.

To explain it, we begin with a simple hypothetical economy where there are only two sectors, the household and the business. The economic activity takes the form of flow of goods and services between these two sectors and monetary flow between them. These two flows, called real and monetary are shown in figure.



Where the product market is shown in the upper portion and the factor market in the lower portion. In the product market, consumers (Household) purchase goods and services from producers (Firms) while in the factor market, consumers receive income from the former for providing factor services. The producers, in turn, make payments to consumers for the services rendered i.e. wage payments for labour services, interest

for capital supplied, etc. thus payments go around in a circular manner from producers to consumers and from consumers to producers, as shown by arrows in the inner portion of the figure. There are also flows of goods and services in the opposite direction to the money payments flows. Goods flow from the business sector to the household sector in the product market, and services flow from the household sector to the business sector in the factor market, as shown in the outer portion of the figure. These two flows are linked by product prices and factor prices. The economy is in general equilibrium when a set of prices is allowed at which the magnitude of income flow from producers to consumers is equal to the magnitude of the money expenditure from consumers to producers.

An example will make the concept of general equilibrium clearer. Suppose the demand for India-manufactured consumer goods suddenly increases in Western Europe. Indian exports will increase thereby increasing output, employment and profits in the export industries. Resources will be diverted from other industries to the export industries.

The demand and prices of the substitute commodities will also increase. The increased demand for exports will have economy-wide effects. An all-round analysis of the repercussions of the economic disturbance increased demand for manufactured consumer goods for export can be done only through general equilibrium theory.

General equilibrium analysis deals with the equilibrium of the whole organisation in the economy: consumers, producers, resource-owners, firms and industries. Not only should individual consumers and firms be in equilibrium in themselves but also in relation to one another.

Business firms enter product markets as suppliers, but they enter factor markets as buyers. Households, on the other hand, are buyers in product markets but suppliers in factor markets. General equilibrium prevails when both the product and factor markets are in equilibrium in relation to each other.

The above review of the utility – maximizing behaviour of the household, and the profit-maximizing behaviour of the firm was conducted under the assumption of competitive conditions (in which commodity prices, income, state of technology, and factor prices are given) and in complete isolation from the influences arising from other parts of the economy. Similarly, the analysis of the product markets, where buyers and sellers interact will each other and among themselves to determine prices and quantities of various factors employed is conducted on the basis of the *ceteris-paribus* assumption. This ignores the interrelationship among the various commodities and the factor markets. In each case the analytical approach is that of partial equilibrium.

However, a fundamental feature of any economic system is the inter-dependence among the constituent parts. The markets of all commodities and all productive factors are inter-related and the prices in all markets are simultaneously determined.

The interdependence between individuals and markets requires that equilibrium for all product and factor markets as well as for all participants in each market must be determined simultaneously in order to secure a consistent set of prices. General equilibrium emerges from the solution of a simultaneous equation model, of millions of equations in millions of unknowns. The unknowns are the prices of all factors and all commodities, and the quantities purchased and sold (of factors and commodities)

### 3.2.1 Assumptions

- (1) A perfectly competitive market system exists both in the product and the factor markets so that consumers and producers face the same set of prices ( $P_x, P_y, w, r$ ).
- (2) The two factors of production are labour (L) and capital (K) whose quantities are exogenously given and which are homogeneous and perfectly divisible.
- (3) Only two commodities, x and y, are produced. Technology is given and the production functions of x and y are represented by smooth and convex to the origin isoquants – implying diminishing marginal rate of factor (technical) substitution along any isoquant. Each production function exhibits constant returns to scale. It is further assumed that the two production functions are independent.
- (4) There are two consumers in the economy, A and B, whose preferences are represented by ordinal indifference curves which are convex to the origin, exhibiting diminishing marginal rate of substitution between the two goods. It is further assumed that the choices of the two consumers are independent. Finally, it is assumed that the choices of the consumers are not influenced by advertising or other activities of the firms.
- (5) The goal of each consumer is to maximize his satisfaction subject to his income constraint, similarly, each producer is assumed to seek to maximize his profit subject to the technological constraint of the production function.
- (6) The factors of production are owned by the consumers
- (7) There is full employment of the factors of production and income received by their owners (A and B) are spent.

Given the above assumptions, a general equilibrium is reached when the four markets (two commodity markets and two factor markets) are cleared at a set of equilibrium prices ( $P_x, P_y, w, r$ ) and each participating economic agent (2 firms and 2 consumers) is simultaneously in equilibrium. The general equilibrium solution thus requires the determination of the values of the following variables:

- (a) The total quantities of x and y, which will be produced by firms and bought by consumers.
- (b) The allocation of K and L to the production of each commodity ( $K_x, K_y, L_x, L_y$ ).
- (c) The quantities of x and y which will be bought by the two consumers ( $X_A, Y_A, X_B, Y_B$ )
- (d) The prices of commodities ( $P_x$  and  $P_y$ ) and of the factors of production (wage  $w$ , and rental of capital,  $r$ )
- (e) The distribution of factor ownership between the two consumers ( $K_A, K_B, L_A, L_B$ )

The quantities of factors multiplied by their prices define the income distribution between A and B, and hence their budget constraint.

### 3.2.2 Static Properties of a General Equilibrium State

Three static properties are observed in a general equilibrium solution, reached with a free competitive market mechanism:

- (a) Efficient allocation of resources among firms (equilibrium of production)
- (b) Efficient distribution of commodities produced between the two consumers (equilibrium of consumption)
- (c) Efficient combination of products (simultaneous equilibrium of production and consumption).

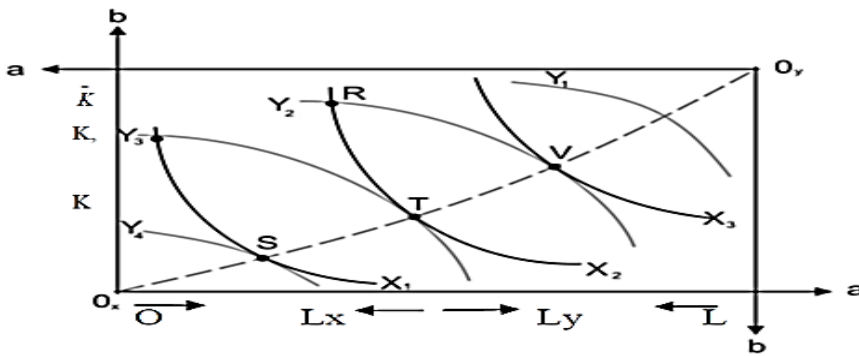
These conditions are called marginal conditions of Pareto optimality.

- (a) Equilibrium of production (Efficiency in Factor Substitution): We saw in Fig. 3.2 that the equilibrium of the firm requires that

$$MRTS_{L,K} = \frac{w}{r} \dots \dots \dots (3.3)$$

where w and r are the factor prices prevailing in the market and MRTS is the marginal rate of technical substitution between the two factors.

The joint equilibrium of production of the two firms in our simple model can be derived using the Edgeworth box of production.



**Figure 3.2: Edgeworth Box of production**

The locus of points of tangency of the x and y isoquants is called the Edgeworth contract curve of production and is given by ox-oy in Figure 3.2. The curve contains the efficient allocations of K and L between the firms. Each point shows a specific allocation of K and L in the production of commodities x and Y that is efficient. Since the Edgeworth contract curve is a locus of tangencies of X and Y isoquants, at each one of its points the slopes of the isoquants are equal:

That is: slope of x isoquant = slope of y isoquant or  $MRTS_{LK}^x = MRTS_{L,K}^y \dots \dots \dots (3.4)$

Thus, in our simple general equilibrium model, the firms will be in equilibrium only if they produce somewhere on the Edgeworth contract curve. This follows from the fact that the factor prices facing the producers are the same and the profit maximization requires that each firm equates its  $MRTS_{L,K}$  with the ratio of factor prices  $\frac{w}{r}$

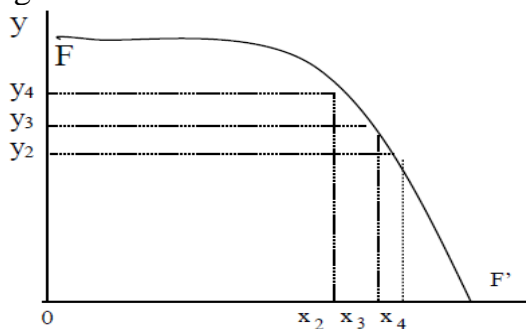
$$MRTS_{LK}^x = MRTS_{L,K}^y = \frac{w}{r} \dots \dots \dots (3.5)$$

This will be the same for all firms.

It thus appears that the production equilibrium is not unique since it may occur at any point along the Edgeworth contract curve. However, with perfect competition (in which all producers face the same factor prices) one of these equilibria will be realized: that is

the one at which the (equalized between firms)  $MRT_{L,K}$  is equal to the ratio of factor prices  $\frac{w}{r}$ . That is, with perfect competition general equilibrium of production occurs where condition (3.5) is satisfied.

If factor prices are given, from the Edgeworth box of production we can determine the amounts of  $x$  and  $y$  which maximize the profits of firms. However, in a general equilibrium analysis, these quantities of  $x$  and  $y$  must be equal to what the consumers want to buy in order to maximize utility or satisfaction. In order to bring together the production side of the system with the demand side, we must define the equilibrium of the firms in the product space, using the production possibility curve or frontier of the economy as shown in Figure 3.3.



**Figure 3.3: Production possibility curve.**

This is derived from the Edgeworth contract curve of production, by mapping its points on a graph on whose axes are measured the quantities of the final commodities  $x$  and  $y$ . From each point of the Edgeworth contract curve of production, we can read off the maximum obtainable quantity of one commodity, given the quantity of the other, with the given factors  $\bar{K}$  and  $\bar{L}$ .

The production possibility curve of an economy is the locus of all Pareto-efficient outputs, given the resource endowment ( $\bar{K}$  and  $\bar{L}$ ) and the state of technology. It is also called the product transformation curve because it shows how a commodity is transformed into another, by transferring some factors from the production of one commodity to the other. Its negative slope is called the marginal rate of product transformation,  $MRPT_{x,y}$ , and, it shows the amount of  $y$  that must be sacrificed in order to obtain an additional unit of  $x$ . by definition

$$MRPT_{x,y} = \frac{dy}{dx} \dots \dots \dots (3.6)$$

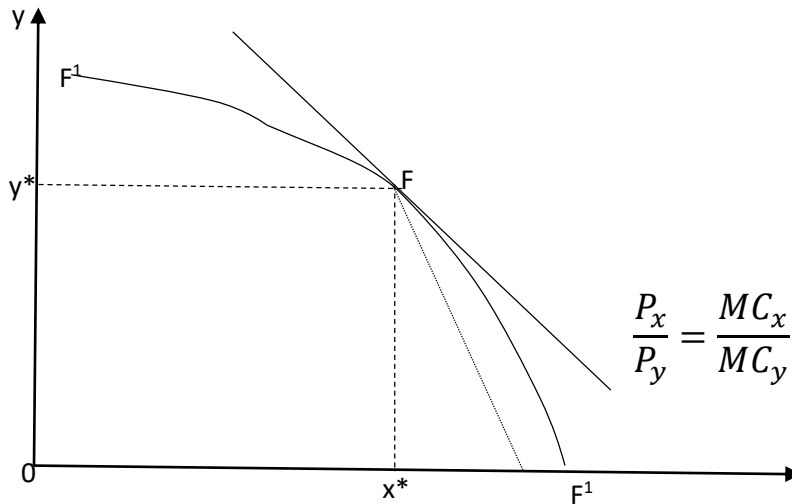
It can be shown that  $MRPT_{x,y}$  is equal to the ratio of the marginal costs of the two products.

$$MRPT_{x,y} = \frac{dy}{dx} = \frac{MC_x}{MC_y} \dots \dots \dots (3.7)$$

In perfect competition, the profit maximizing firm equates the price of the commodity produced to the long-run marginal cost of production:

$$MC_x = P_x \text{ and } MC_y = P_y \dots \dots \dots (3.8)$$

Thus given the commodity price, the equilibrium is reached at the point on the production transformation curve that has a slope equal to the ratio of these prices, as shown in Figure 3.4.



**Figure 3.4: General Equilibrium of production in perfect competition.**

At this point, the product mix is given by  $x^*$  and  $y^*$

Proof of Equation. 3.7:

By definition  $MC_x = d \frac{TC}{dx}$  and  $MC_y = d \frac{TC}{dy}$  ..... (1)

Thus  $\frac{MC_x}{MC_y} = d \frac{(TC)x}{dx} \times \frac{dy}{d(TC)y} = \frac{d(TC)x}{d(TC)y} \frac{dy}{dx}$ .....(2)

But  $d(TC)x = w(dL_x) + r.(dK_x)$

And  $d(TC)y = w (dL_y) + r.(dK_y)$

So that  $\frac{d(TC)x}{d(TC)y} = \frac{(w(dL_x)+r.(dK_x))}{(w (dL_y)+r.(dK_y))}$  ..... (3)

In order to remain on the production possibility curve (PPC) the factors released from the decrease in commodity Y must be equal to the factors absorbed by the increase in the production of x, that is

$dl_x = -dl_y$  and  $dk_x = -dk_y$  ..... (4)

Substituting (4) in (3) we get:

$$\frac{d(TC)x}{d(TC)y} = \frac{(w(-dL_y) + r.(-dK_y))}{(w (dL_y) + r.(dK_y))} = -1$$

and substituting (5) in (2) we obtain

$$\begin{aligned} \frac{MC_x}{MC_y} &= -1 \left( \frac{dx}{dy} \right) = \frac{dy}{dx} = \text{slop of PPC} \\ &= MRPT_{x,y} \dots \dots \dots QED \dots \dots \dots (6) \end{aligned}$$

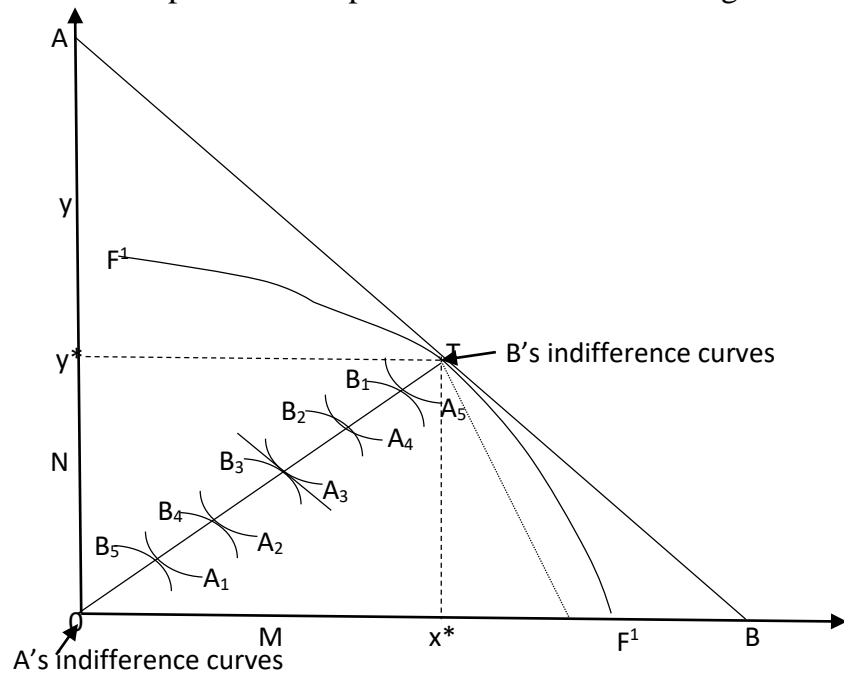


**Equilibrium of Consumption (Efficiency in Distribution of Commodities)**

From Figure 3.1 we saw that the consumer is in equilibrium when  $MRS_{x,y} = \frac{P_x}{P_y}$ . Since both consumers in a perfectly competitive market are faced with the same prices, the condition for joint or general equilibrium of the consumer is:

$$MRS_{xy}^A = MRS_{xy}^B = \frac{P_x}{P_y} \dots \dots \dots (3.9)$$

This general equilibrium of consumption for the product mix is shown in Fig 3.6.



**Figure 3.5: Equilibrium of Consumption (in a perfect market)**

In the figure above, the Edgeworth contract curve of consumption, OT, represents efficient distributions. At each point on this curve, the following equilibrium condition is satisfied

$$MRS_{x,y}^A = MRS_{x,y}^B \dots \dots \dots (3.10)$$

Thus, for a given product mix such as T, there is an infinite number of possible efficient or optimal equilibrium distributions (indicating that equilibrium of consumption is not unique). However, with perfect competition, only one of these points is consistent with the general equilibrium of the system. This is the point of the contract curve where the ('equalized')  $MRS_{x,y}$  of the consumers is equal to the price ratio of the commodities, that is, the condition (3.9) is satisfied.

In the figure, the equilibrium of the consumers is defined by point, T. Consumer A reaches the utility level implied by the indifference curve  $A_3$ , buying OM of x and ON of y. Consumer B reaches the utility level implied by the indifference curve  $B_3$ . Buying  $MX_B$  of x and  $Ny_0^s$  of Y.

**(b) Simultaneous Equilibrium of Production and Consumption (Efficiency in Product – Mix)**

In addition to the conditions specified in (a) and (b) above, a third condition is required to be satisfied if the general equilibrium of the system as a whole is to be achieved. This

condition is that the marginal rate of product transformation (slope of the PPC) be equal to the marginal rate of substitution of the two commodities between the consumers.

$$MRPT_{xy} = MRS_{x,y}^A = MRS_{x,y}^B \dots \dots \dots 3.11)$$

In perfect competition, this condition is satisfied since, from equation (3.8)

$$MRPT_{x,y} = \frac{P_x}{P_y} \dots \dots \dots (3.12)$$

and from equation (3.9)  $MRS_{x,y}^A = MRS_{x,y}^B \dots \dots \dots (3.13)$

so that:

$$MRPT_{x,y} = MRS_{x,y}^A = MRS_{x,y}^B \dots \dots \dots (3.14)$$

(as demonstrated by the equality of slopes at T and T in Fig. 3.6). This is the third marginal condition of Pareto efficiency. It refers to the efficiency of product substitution or optimal composition of output.

In summary, with perfect competition, the simple two factors, two-commodities, two consumer systems have a general equilibrium solution in which three Pareto efficiency conditions are satisfied.

1. The MRS between the two goods is equal for both consumers (optimal allocation of the goods among consumers)
2. The MRTS between the two factors is equal for all firms (efficiency in factor substitution-implies optimal allocation of factors among the two firms)
3. The MRS and the MRPT are equal for the two goods (optimal composition of output and thus optimal allocation of resources).

**SELF ASSESSMENT EXERCISES**

1. What do you understand by the term general equilibrium?
2. In an economy of two individuals (A and B) and two commodities (X and Y), state the condition(s) at which general equilibrium of exchange is reached.

**3.3 The Walrasian System: Algebraic Approach**

Leon Walras, a French economist, in his book “Elements of Pure Economics” argued that all prices and quantities in all markets are determined simultaneously through their interaction with one another. He used a system of simultaneous equations to describe the interaction of individual sellers and buyers in all markets, and maintained that all the relevant magnitudes (prices and quantities of all commodities and all factor services) can be determined simultaneously by the solution of this system.

In the Walrasian Model, the behaviour of each individual decision-making is presented by a set of equations. For instance, each consumer has a double role: he buys commodities and sells services of factors to firms. Thus for each consumer we have a set of equations consisting of two subsets: one for the demands for different commodities and the other for his supplies of factor inputs. Similarly, the behaviour of each firm is presented by a set of equation with two subsets: one for the quantities of commodities that it produces, and the other for the demand for factor inputs for each commodity produced. The important characteristic of these equations is their

simultaneity or interdependence. The solution of this system of millions of simultaneous equations defines the ‘unknowns’ of the model. The unknowns are the prices and quantities of all commodities and all factor inputs.

In a general equilibrium of the Walrasian type, there are as many markets as there are commodities and factors of production.

For each market there are three types of functions: demand functions, supply functions and a function which clears the market by stipulating that quantities demanded are equal to the quantities supplied.

In a commodity market, the number of demand functions is equal to the number of consumers, and the number of the supply functions is equal to the number of firms which produce the commodity.

In each factor market the number of demand functions is equal to the number of firms multiplied by the number of commodities they produce. The number of supply functions is equal to the number of consumers who own the factors of production.

A necessary, but not sufficient, condition for the existence of a general equilibrium is that there should be in the system as many independent equations as there are number of unknowns.

For example, assume that an economy consists of two consumers, A and B, who own two factors of production, L and K. These factors are used by two firms to produce two commodities, x and y. (This is a simple  $2 \times 2 \times 2$  general equilibrium model). It is assumed that each firm produces one commodity, and each consumer buys some quantity of both.

It is also assumed that both consumers own some quantity of both factors, the distribution of which ownership is exogenously determined. In this simple model we have the following.

i.	Quantities demanded of x and y by consumers $2 \times 2 =$	4
ii.	Quantities supplied of K and L by consumers $2 \times 2 =$	4
iii.	Quantities demanded of K and L by firms $2 \times 2 =$	4
iv.	Quantities of y and x supplied by firms	2
v.	Prices of commodities y and x	2
vi.	Prices of factor L and K	2
<b>vii.</b>	<b>Total number of ‘unknowns’</b>	<b>18</b>

To find these unknowns we have the following number of equations:

(i)	Demand functions of consumers $2 \times 2 =$	4
(ii)	Supply functions of factors $2 \times 2 =$	4
(iii)	Demand functions for factors $2 \times 2 =$	4
(iv)	Supply functions of commodities	2
(v)	Clearing-the-market of commodities	2
(vi)	Clearing-the-market of factors	2
<b>(vii)</b>	<b>Total number of equation</b>	<b>18</b>

Since the number of equations is equal to the number of unknowns, one is tempted to think that a general equilibrium solution exists. Unfortunately, this equality is neither a necessary nor a sufficient condition for the existence of a solution. In the Walrasian system, one of the equations is not independent of the other: there is a ‘redundant

equation' in the system which deprives the system of a solution, since the number of unknowns is larger than the number of independent equations. In this model, the absolute level of prices cannot be determined. General equilibrium theorists have adopted the device of choosing arbitrarily, the price of one commodity as a numeraire (or unit of account) and express all other prices in terms of the price of the numeraire. If we assign unity to the price of the numeraire, we attain equality of the number of simultaneous equations and unknown variables (the number of unknown is reduced to 17 in our example). However, the absolute prices are still not determined. They are simply expressed in terms of the numeraire. This indeterminacy can, however, be *criminated* by the introduction of the money market in the model.

### 3.3.1 Equations of the Model

The general equilibrium model can be extended to any number of households' commodities, and factors of production. For example, we may let the economy consist of  $H$  household,  $M$  commodities and  $N$  types of inputs. This is the  $H \times M \times N$  general equilibrium model.

In the model, the subscript  $h$  ( $h = 1, 2, \dots, H$ ) is used to denote a particular household. The subscript  $m$  ( $m = 1, 2, \dots, M$ ) is used to denote a particular commodity (and the corresponding producing firm). Finally, the subscript  $n$  ( $n = 1, 2, \dots, N$ ) is used to denote a particular factor of production.

Let  $q_m$  denote the quantities of commodities: the symbol  $V_n$  denote the quantities of factors of production. The prices of commodities are denoted by the symbol  $p_m$  and the prices of factors,  $W_n$ . All the assumptions of the  $2 \times 2 \times 2$  model hold for the general model. In particular, there is perfect competition in all markets, so that consumers and firms are faced by the same set of prices.

#### A. The Household Sector

The demand function for the  $m^{\text{th}}$  commodity by the  $h^{\text{th}}$  consumer is of the general form.

$$q_{mh} = f_{mh}(P_1, P_2, \dots, P_M, W_1, W_2, \dots, W_N) \dots \dots \dots (4.1)$$

$$(m = 1, 2, \dots, M)$$

$$(h = 1, 2, \dots, H)$$

(tastes are implicit in the  $mh$  symbol of the right-hand side).

There are  $MH$  demand functions in the general system. The supply function for the  $n^{\text{th}}$  factor by the  $h^{\text{th}}$  consumer is of the general form.

$$V_{nh} = f_{nh}(P_1, P_2, \dots, P_M, W_1, W_2, \dots, W_N) \dots \dots \dots (4.2)$$

$$(n = 1, 2, \dots, N)$$

$$(h = 1, 2, \dots, H)$$

(Tastes are again implicit in the  $f_{nh}$  symbol).

There are  $NH$  supply functions for the productive services

#### B. The Business Sector

The demand for the  $n^{\text{th}}$  factor to be used in the production of the  $m^{\text{th}}$  commodity depends on the quantity of this particular commodity, the prices of all factors of production and the production function.

$$V_{nm} = f_{nm}(q_m, W_1, W_2, \dots, W_N) \dots \dots \dots (4.3)$$

$$(M = 1, 2, \dots, M)$$

$$(N = 1, 2, \dots, N)$$

(The particular production function of the firm is implicit in the  $f_{nm}$  symbol. It is assumed that the production isoquants are smooth and convex to the origin, and there are constant returns to scale).

There are  $NM$  demand equations for productive services. The supply of any commodity in perfect completion is defined, in the long run, by the equality of the revenue to the long run average cost.

$$P_m q_m = \sum_{n=1}^N V_{nm} W_n \dots \dots \dots (4.4)$$

$$(m = 1, 2, \dots, M)$$

$$(n = 1, 2, \dots, N)$$

There are  $M$  supply equations for the  $M$  commodities produced in the economy

### C. Equilibrium Conditions (or Clearing of the Market Equations)

In order to have equilibrium in all markets the quantity supplied must be equal to the quantity demanded, or equivalently, the excess demand must be zero in each market.

Equilibrium in each commodity market requires that the total demand of households for that product to equal total supply by firms of that product. Hence we have  $M$  “clearing of the commodity market” equations of the form.

$$\sum_{h=1}^H W_a q_{mh} = q_m (m = 1, 2, \dots, M) \dots \dots \dots (4.5)$$

Similarly, for each factor market to be in equilibrium the total demand by firms for that factor must be equal to the total supply by households of that factor. Thus we have  $N$  “clearing of the factor market” equations of the form.

$$\sum_{m=1}^M V_{nm} = \sum_{h=1}^H V_{nh} (n = 1, 2, \dots, N) \dots \dots \dots (4.6)$$

In total the  $H \times M \times N$  general model consists of the following equations.

### I. Behavioural Equations

- i. Demand functions for commodities      MH
- ii. Supply functions for commodities      M
- iii. Demand function for factors      NM
- iv. Supply functions for factors      NH

### II. Clearing the Market' Equations

- |     |                                   |   |
|-----|-----------------------------------|---|
| v.  | Clearing of the commodity markets | M |
| vi. | Clearing of the factor markets    | N |

$$\text{Total} = MH + M + NM + NH + M + N$$

#### D. Number of Unknowns

The unknowns that must be determined by the solution of the general model are:

- The quantities of goods demanded by the households,  $q_{mh}$ . There are  $MH$  such quantities.
- The quantities of goods supplied by the firms,  $q_{mh}$ . There are  $M$  such variables.
- The quantities of productive factors demanded by the firms,  $V_{nm}$ . There are  $NM$  such quantities.
- The quantities of productive factors supplied by the households,  $V_{nh}$ . There are  $NH$  such quantities.
- The prices of commodities that clear their markets,  $P_m$ . There are  $M$  such prices.
- The prices of factors that clear their markets,  $W_n$ .

In total the number of unknowns is  $MH + M + NM + NH + M + N$

However, the independent equations are one short of the number of unknowns. There is one redundant (not independent) equation. Thus the entire system is undermined, and we cannot find unique values for all commodity and factor prices and for all commodity and factor quantities.

The indeterminacy can be resolved by choosing arbitrarily any commodity or factor as the numeraire and expressing the prices of all other commodities and factors in terms of its price (which for convenience is set equal to 1, just as the price of  $N_1$  is 1). In this way there is one less unknown to be determined by the solution of the model.

The system with the numeraire is called a real model, because prices are stated in terms of a commodity instead of being expressed in terms of money. The above general equilibrium system of equations determines unequally the price ratios, but not the absolute level of prices. However, all quantities of commodities and factors are unequally determined. The absolute level of prices can only be determined if the real general equilibrium model is augmented to include a market for money.

#### SELF ASSESSMENT EXERCISE

List the Walrasian assumptions of general equilibrium

### 3.4 Dynamic Equilibrium

Dynamic equilibrium is said to exist when after a fixed period, the equilibrium state is disturbed. In dynamic equilibrium, prices, quantities, incomes, tastes, technology etc are constantly changing.

For e.g. suppose some more persons develop the taste for fish, as a result of this, the demand for fish will increase and the sellers will automatically raise the price and thus change the behaviour of the old buyers. The market will be thrown into a state of disequilibrium and will remain so until the supply of fish is increased to catch up with

the level of the new demand. Then, a new equilibrium will be brought in by the forces contending forces.

The word dynamic, simply means causing to move. In economics, 'dynamic' refers to the study of economic change. The essence of any knowledge lies in formulating relationships between phenomena. There must be thus sequence of events for the knowledge to be born. The main purpose is to explore how current events will shape themselves in the future. To do so, it is necessary to visualize the way they arose from out of the past events. The moment we talk of sequence of events, the elements of time creeps into our analysis. Dynamic equilibrium, in the language of economics, therefore tries to explain the process of change through time.

Dynamic equilibrium is of two types

1. Micro Dynamic equilibrium
2. Macro Dynamic equilibrium

### **Micro-Dynamic (cobweb) equilibrium**

This is used to explain the changes of demand, supply and price over a long period of time. The cob-web model (or theorem) analyses the movements of prices and outputs when supply is wholly determined by prices in the previous period.

As prices moves up and down in cycles, quantities produced also seem to move up and down in a counter-cyclical manner (e.g. prices of perishable commodities like vegetables).

In order to find out the conditions for converging, diverging or constant cycles: one has to look at the slope of the demand curve and then of the supply curve.

### **Assumption**

The cob-web model is based on the following assumption:

1. The current year's (t) supply depends on the last year's ( $t-1$ ) decisions regarding output level.
2. Hence current output is influenced by last year's price. i.e.  $P(t-1)$
3. The current period or year is divided into sub-periods of a week or fortnight.
4. The parameters determining the supply function have constant values over a series of periods.
5. Current demand ( $D_t$ ) for the commodity is a function of current price ( $P_t$ ).
6. The price expected to rule in the current period is the actual price in the last year.
7. The commodity under consideration is perishable and can be stored only for one year.
8. Both supply and demand function are linear .i.e. both are straight line curves which increase or decrease at a constant proportion.

### **The Cob-web Model**

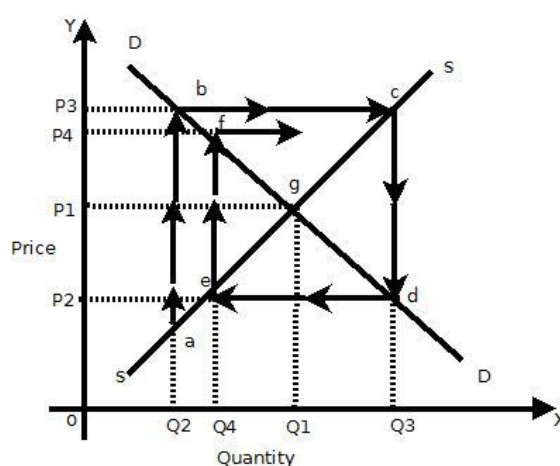
There are three types of Cob-web Models:

1. **Convergent**
2. **Divergent**
3. **Continuous**

### (1) **Convergent Cob-web**

Under this model, the supply is a function of previous year i.e.  $S = f(P_{t-1})$  ('t' is the current period and 't-1' is a previous period) and on the other hand the demand is the function of price i.e.  $D_t = f(P)$ . The equality between the quantity supplied and quantity demanded is called market equilibrium. i.e.  $S_t = D_t$ . Equilibrium can be established only through a series of adjustment if current supply is in response to the price during the last year. But this adjustment will take place over several consecutive periods.

For e.g. let us assume onion growers produce only one crop in a year. The onions grown this year will be based on the assumption that the price of onions will equal that of last year. The market demand and supply curves for onions are represented by DD and SS curves respectively in diagram below.



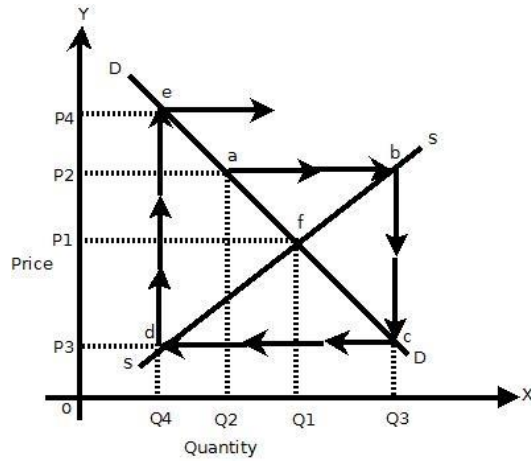
**Figure 3.6: Convergent Cob-web**

Suppose the price of last was  $OP_1$  and Producers decide the equilibrium output  $OQ_1$  this year. Now suppose there is crop failure due to natural calamities which decrease the output  $OQ_2$  which is less than  $OQ_1$  (i.e. equilibrium output). This shortage will increase the price to  $OP_2$  in the current period. In the next period, the onion growers will produce  $OQ_3$  quantity in response to the higher price  $OP_2$ . This will be more than the equilibrium quantity  $OQ_1$  which is the actual need of the market. The excess supply will lower the price to  $OP_3$ . This will encourage the producer to change the production plan. They will reduce the supply to  $OQ_4$  in the third period. But this quantity is less than the equilibrium quantity  $OQ_1$ . This will lead to again rise in price to  $OP_4$ , which in turn will encourage the producers to produce  $OQ_1$  quantity. The equilibrium will be established at point g where DD and SS curves intersect. This series of adjustments from point a, b, c, d, e to f is traced out as a cobweb pattern which converge towards the point of market equilibrium g. This is also called dynamic equilibrium with lagged adjustment.

### (2) **Divergent Cob-Web**

The divergent cob-web is unstable when price and quantity changes move away from the equilibrium position. This can be explained with the help of following diagram.





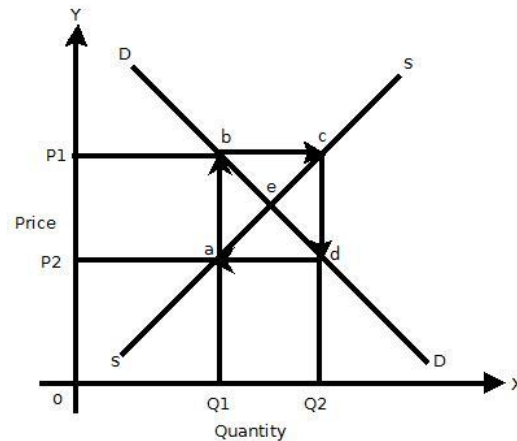
**Figure 3.7: Divergent Cob-web**

We will start with the initial equilibrium price  $OP_1$  and equilibrium quantity  $OQ_1$ . Now suppose there is a temporary disturbance that causes output to fall to  $OQ_2$ . Due to lack of supply, the price will rise to  $OP_2$ .

The increase in price will in turn raise the output to  $OQ_3$  which is more than the equilibrium output  $OQ_1$ . The increase in supply will lead to fall in price to  $OP_3$ . This fall in price will increase the demand and there will be excess demand  $OQ_2$  than supply. The excess demand will shoot up the price to  $OP_4$ . This shows that the price will still go away from the equilibrium after the adjustment by the producers. This is called a divergent cob-web.

**(3) Continuous cob-web**

The cob-web models show the continuous changes in price and quantities.



**Figure 3.8: Continuous Cob-web**

Suppose we start with the price  $OP_1$  as shown in the diagram. The supply will be more than the demand due to the high price. On the other hand, the demand will be  $OQ_2$  which is less than the supply. This fall in demand will force the producer to decrease price to  $OP_2$  in next period. But at this price  $OP_2$  the demand will be  $OQ_2$  which is more than the supply  $OQ_1$ . In this way, the prices and quantities will circulate constantly around the equilibrium.

## Macro-Dynamics (cobweb)

According to Kurihara, 'Macro-Dynamics' treats discrete movements or rates of change of macro-variables. It can be explained in terms of the Keynesian process of income propagation (the investment multiplier) where consumption depends on income i.e.  $C = f(Y_{t-1})$

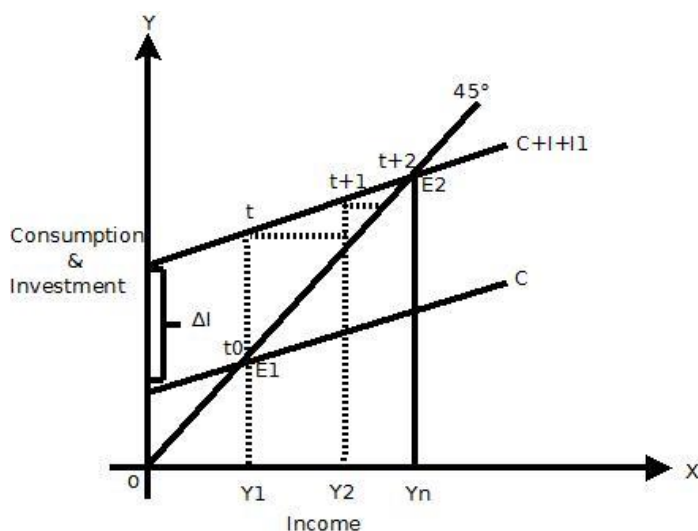
**Where:**

**C = Consumption**

**Y = Income**

**f = function**

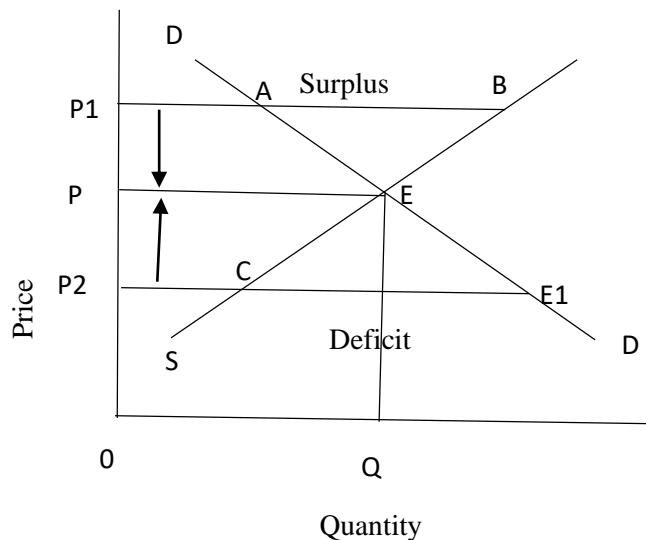
The function shows that the consumption in the current period (t) depends on the Y in the previous period ( $t-1$ ). On the other hand, investment is a function of time and of constant autonomous investment  $\Delta I$  (Autonomous investment is the investment which does not change due to changes in income. i.e. changes in investment does not take place due to change in income). For e.g. government does the investment for welfare of the people and not for profit expectation. So investment function can be written as  $I_t = f(\Delta I)$ . This can be explained with the following diagram.



The 45-degree line is the aggregate supply function. Suppose we start with the time period  $t_0$  where with an equilibrium level of income  $OY_1$ , investment increased from  $I_0$  to  $I_1$ , this can be seen by the new aggregate demand function line  $C + I + I_1$ . But in period  $t$ , consumption lags behind and it is still on the equilibrium point  $E_1$ . In next period,  $(t + 1)$ , consumption will increase with the increase in investment. This also increases income from  $OY_1$  to  $OY_2$ . This is the process of income propagation which will continue till the aggregate demand function  $C + I + I_1$  intersects the aggregate function  $45^\circ$  line at point  $E_2$  in the  $n^{\text{th}}$  period. The new equilibrium level of income is at  $OY_n$ . The curved steps from  $t_0$  to  $E_2$  show the macro-dynamic equilibrium path.

### (1) Stable Equilibrium

Stable equilibrium exists when an economic unit or system has the tendency to regain equilibrium as soon as it is disturbed. This means that in the event of any disruption, there is tendency for the initial equilibrium to be restored. In other words, it is self-adjusting. This is because there are economic forces of correction that readily push the system back to its original position. Stable equilibrium can be illustrated using the graph below.



**Figure 3.9: Stable Equilibrium**

In figure above, DD is demand curve which is negatively sloped and SS is the supply curve which is positively sloped. The equilibrium occurs at point E where the supply and demand are in balance and equal. The equilibrium price OP and the equilibrium quantity OQ are determined. This is an example of a stable equilibrium in economics.

Assuming that the market price is OP1. At this price, P1B is the quantity supplied with P1A being the quantity demanded. This shows that quantity supplied is greater than quantity demanded with a surplus quantity represented by AB. This leads to a downward pressure on price. The downward pressure applies until the price reaches the equilibrium level at which the quantity supplied equals the quantity demanded.

Also, at the price level OP2, the quantity supplied is less than the quantity demanded with CE1 showing the amount of shortage of commodity. Thus there will be excess demand which will lead to an upward pressure on the price that eventually pushes up the price to the equilibrium level at which the quantity supplied equals the quantity demanded.

## (2) Unstable Equilibrium

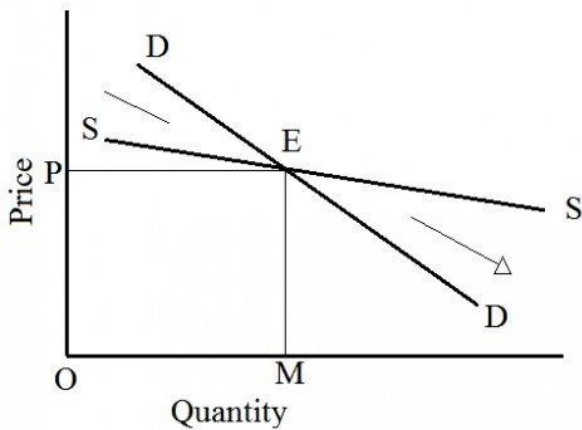
Equilibrium is unstable when as soon as the initial position is disturbed, it generates further chain of disturbance that pushes the equilibrium point further away with no tendency for any restoration. Thus, in case of unstable equilibrium, the disturbance in the economy will lead or exaggerate further disturbances and will never regain the original position. In supply and demand analysis, unstable equilibrium can occur at two occasions:

- when there is a negatively sloped supply curve and
- when there is a positively sloped demand curve

### When there is a negatively sloped supply curve

In this case, equilibrium is unstable when there is an interaction between a normal negatively sloping demand curve and negatively sloped supply curve, which is a rare and exceptional case. This negatively sloping supply curve is possible when both increasing production and decreasing costs occur simultaneously due to various internal and external economies of scale enjoyed by the firm.

Figure 3

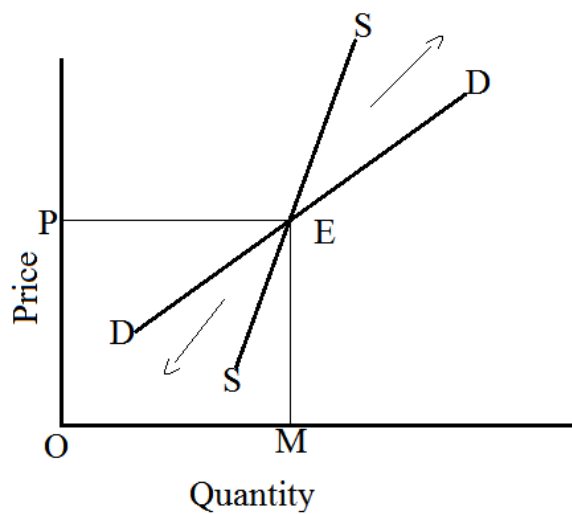


**Figure 3.10: Unstable Equilibrium I**

In the figure, point E is the equilibrium, OP is the equilibrium price and OM is the equilibrium quantity. When price rises above equilibrium price, the quantity demanded becomes greater than the quantity supplied. This leads to a rise in price due to the excess demand, thereby moving away from equilibrium. Similarly, when the price falls below equilibrium, the quantity supplied is greater than the quantity demanded. Since there is excess supply, the price falls further and continues to move away from the equilibrium. In both cases, there is no possibility for the price to move towards equilibrium. Hence, E represents an unstable equilibrium position.

**When there is a positively sloped demand curve**

Here, equilibrium is unstable when there is an interaction between an abnormal positively sloping demand curve and a normal positively sloped supply curve, which is seen with Giffen goods.

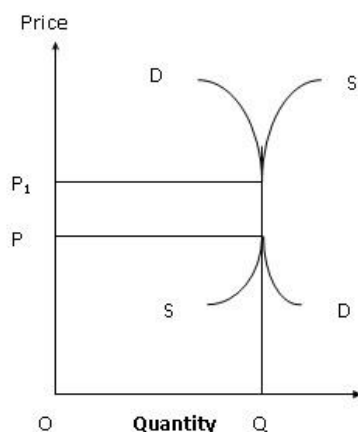


**Figure 3.11: Unstable Equilibrium II**

In the figure, the demand curve intersects the supply curve at E with the equilibrium price being OP, while equilibrium quantity is OM. An increase in price above OP leads to an excessive amount of demand over supply. This causes an increase in the price. A reduction in price below OP contributes to excess supply over demand. This excess supply over demand triggers a further decrease in the price. Hence, E in the above diagram is in unstable equilibrium since there is no chance for the original equilibrium to be restored

### (3) Neutral Equilibrium

Neutral equilibrium is when the disturbing forces neither bring it back to the original position nor do they drive it further away from it. It rests where it has been moved. When an initial equilibrium position is disturbed, the forces of disturbance bring it to the new position of equilibrium where the system has come to rest. The situation of neutral equilibrium arises when demand and supply curves go together in a range of prices or in a range of quantities. The static neutral equilibrium condition is illustrated in figure.



**Figure 3.11: Neutral Equilibrium**

Figure 3.11 is the initial equilibrium point where OQ quantity is demanded and supplied at OP price. With the rise in the price to OP1, E1 becomes the new equilibrium point but the quantity demanded and supplied remains the same, i.e. OQ. Thus, the price range PP1 (=EE1) represents neutral equilibrium.

## 4.0 CONCLUSION

General equilibrium analysis is an extensive study of a number of economic variables and elements, their interrelations and interdependences for understanding the working of the economic systems as a whole. Partial equilibrium is the study of the equilibrium position of an individual, a firm, an industry or a group of industries viewed in seclusion. General equilibrium considers how the economy will be efficient if for any amount of X produced under cost minimization, the remaining resources are employed in a way that will yield maximum possible amount of Y. On the contract curve, it is impossible (inefficient) to increase production of one of the two goods without decreasing that of the other. Partial equilibrium analysis of markets assumes that related

markets are unaffected. General equilibrium analysis examines all markets simultaneously, taking into account feedback effects on other markets being studied.

## **5.0 SUMMARY**

In this unit, we studied the concept of general equilibrium. We went ahead to illustrate in graphical terms how markets are interrelated and attained efficiency. You also learnt about general equilibrium in production and relationship between contract curve and transformation curve/ production possibilities curve. General equilibrium analysis is a system-wide examination of prices and markets, focusing specifically on the interactions within the economic system. 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.

## **6.0 TUTOR-MARKED ASSIGNMENT**

1. How is equilibrium attained in the household and firm?
2. State the major difference between partial equilibrium and general equilibrium.
3. How is general equilibrium attained in two different markets?
4. Explain dynamic equilibrium

## **7.0 REFERENCES /FURTHER READINGS**

- Emery, E.D. (1984). *Intermediate Microeconomics*, Harcourt Brace Jovanovich, Publishers, Orlando, USA
- Jhingan, M.L. (2009). *Microeconomics Theory*, Vrinda Publications (P) Limited, Delhi
- Pindyck, R.S & Daniel, L. R. (2009). *Microeconomics*, (7ed), Pearson Education, Inc. New Jersey, USA.
- Salvatore, D. (2003). *Microeconomics*, Harper Publishers Inc. New York, USA
- Salvatore, D. (2006). *Microeconomics*. Schuams Series. McGraw-Hill Companies, Inc. USA

## **MODULE 2: PRODUCTION FUNCTIONS, MARKET STRUCTURES, DETERMINATION OF WAGES, RENT, INTEREST AND PROFIT**

**Unit 1: Production Function, Duopoly and Oligopoly**

**Unit 2: Bilateral Monopoly and Monopsony**

**Unit 3: Theories of Determination of Wages, Rent, Interest and Profit**

### **UNIT 1: PRODUCTION FUNCTION, DUOPOLY AND OLIGOPOLY**

#### **CONTENTS**

- 1.0 Introduction
- 2.0 Objectives
- 3.0 Main Content
  - 3.1 Production Function
    - 3.1.1 Types of Production Function
  - 3.2 Duopoly
    - 3.2.1 Types of Duopolies
    - 3.2.2 The Significance of a Duopoly
  - 3.3 Oligopoly
    - 3.3.1 Strategic Interactions
    - 3.3.2 Cournot Model
    - 3.3.3 The Bertrand Model
    - 3.3.4 Stackelberg Model
- 4.0 Conclusion
- 5.0 Summary
- 6.0 Tutor-Marked Assignment
- 7.0 References/Further Reading

#### **1.0 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 of production 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.

However, the technical conditions facing the firm are summarised in the production function. The production function is a mathematical statement of the physical relationship given by technology, between a firm's input of productive resources and its output of goods and services per unit of time. 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.

However, a production function is a mirror of a cost function. In other words, as firms maximise their profit, they also minimise their cost.

In this unit we will be looking at duopoly and oligopoly and how production function is determined under this market structures.

## 2.0 OBJECTIVES

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

- ✓ explain the concept of production function, duopoly and oligopoly
- ✓ explain the types of duopoly models
- ✓ explain the types of oligopoly models
- ✓ Solve mathematical questions

## 3.0 MAIN CONTENT

### 3.1 Production Function

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.

Production function is a physical or technical relationship between inputs and output in any given production processes. It describes the rate at which inputs are transformed into outputs. It defines the production possibilities open to the farmers. In an implicit form, the production function is defined as:

$$Q = f(x).$$

It states that output Y is a function of input X.

Where Q = output of product, X = input used,  $f()$  = functional form.

#### 3.1.1 Types of Production Function

Production functions are classified into two based on time period. They are short run and long run production functions.

(a) **Short run production function:** In the short run production function, at least one of the inputs used can be varied with the others fixed. The implicit functional form is defined as:

$$Q = f(X_1, X_2^0).$$

Q = Output

$X_1$  = Labour

$X_2^0$  = Fixed amount of capital

From the short run production function, the following production curves can be defined

**Total Product (TP):** This is the summation of all the outputs produced by varying  $X_1$  holding  $X_2$  constant

$$TP = Q = f(X_1, X_2^0).$$



**Average Product (AP):** This is defined as TP divided by the amount of variable factor employed

$$AP = \frac{TP}{X_1} = f\left(\frac{X_1, X_2^0}{X_1}\right).$$

**Marginal Product (MP):** This is the additional product as a result of employing extra unit of the variable factor  $X_1$ . It can also be defined as the first derivative of the production function with respect to variable factor. It is the slope of the production function

$$MP = \frac{dTP}{dX_1} = \frac{dQ}{dX_1} = f_1(X_1, X_2^0).$$

### The Law of Diminishing Marginal Return

The law states that ‘as more and more of the variable factor is added into the fixed factor, marginal product will increase but beyond an input point, marginal product will fall’. This means that diminishing return occur in the short run when one factor is fixed (capital); and if the variable factor (labour) continues to increase, there comes a point where it will become less productive and therefore there will eventually be a decreasing marginal product.

(b) **Long run production function:** In the long run production function, all the inputs can be varied. It is implicitly expressed as;

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

### Law of Return to Scale

In the long run, there is the law of return to scale which tells us what happens to output as all factors of production changes by the same proportion.

$$Q = f(X_1, X_2).$$

By multiplying  $X_1$  and  $X_2$  by  $R$  (the same proportion shown by  $R$ ) i.e. increasing  $X_1$  and  $X_2$  by  $R$

$$Q^* = f(RX_1, RX_2).$$

If  $Q^*$  increases by the same proportion as increase in inputs, we have constant return to scale. i.e. doubling input leads to doubling output

If  $Q^*$  increases by more than proportionate increase in inputs, we have increasing return to scale. i.e. doubling input leads to more than doubling output

If  $Q^*$  increases by less than proportionate increase in inputs, we have decreasing return to scale. i.e. doubling input leads to less than doubling output.

## 3.2 Duopoly

A duopoly is a type of oligopoly, characterized by two primary corporations operating in a market or industry, producing the same or similar goods and services. The key components of a duopoly are how they affect one another.

In a duopoly, two companies control virtually the entirety of the market for the goods and services they produce and sell. While other companies may operate in the same space, the defining feature of a duopoly is the fact that only two companies are considered major players. These two firms – and their interactions with one another – shape the market they operate in.

### 3.2.1 Types of Duopolies

There are two primary types of duopolies: The Cournot Duopoly (named after Antoine Cournot) and the Bertrand Duopoly (named after Joseph Bertrand).

#### 1. The Cournot Duopoly

Antoine Cournot was a French mathematician and philosopher. In the early to mid-1880s, Cournot used his understanding of mathematics to formulate and publish a significant model of what oligopolies look like. The model, known as the Cournot Duopoly Model (or the Cournot Model), places weight on the quantity of goods and services produced, stating that it is what shapes the competition between the two firms in a duopoly. In Cournot's model, the key players in the duopoly make an arrangement to essentially divide the market in half and share it.

Cournot's model speculates that in a duopoly, each company receives price values on goods and services based on the quantity or availability of the goods and services. The two companies maintain a reactionary relationship in regard to market prices, where each company changes and makes adjustments to their respective production, ending when an equilibrium is reached in the form of equal halves of the market for each firm.

Cournot duopoly, also called Cournot competition, is a model of imperfect competition in which two firms with identical cost functions compete with homogeneous products in a static setting. Cournot's duopoly represented the creation of the study of oligopolies, more particularly duopolies, and expanded the analysis of market structures which, until then, had concentrated on the extremes: perfect competition and monopolies.

Cournot really invented the concept of game theory almost 100 years before John Nash, when he looked at the case of how businesses might behave in a duopoly. There are two firms operating in a limited market. Market production is:  $P(Q) = a - bQ$ , where  $Q = q_1 + q_2$  for two firms. Both companies will receive profits derived from a simultaneous decision made by both on how much to produce, and also based on their cost functions:  $TC_i = C - q_i$ .

So, algebraically:

$$\text{Max } \pi_i(Q) = [a - b(q_i + q_j) - c]q_i$$

In order to maximise, the first order condition will be:

$$q_i = \frac{a - bq_j - c}{2b}$$

And, if  $q_i = q_j$ , then both equal:

$$\frac{a - c}{3b}$$

Therefore, the reaction functions (blue lines), where the key variable is the quantity set by the other firm, will take the following form:

$$RF_i(q_j) = \frac{a - c}{2b} - \frac{1}{2}q_j$$

What all this explains is a very basic principle. Both companies are vying for maximum benefits. These benefits are derived from both maximum sales volume (a larger share of the market) and higher prices (higher profitability). The problem stems from the fact that increasing profitability through higher prices can damage revenue by losing market share. What Cournot's approach does is maximise both market share and profitability by defining optimum prices. This price will be the same for both companies, as otherwise the one with the lower price will obtain full market share, which makes this a Nash equilibrium, also known for this model the Cournot-Nash equilibrium.

If we consider isoprofit curves (those which show the combinations of quantities that will render the same profit to the firm, red curves) we can see that the equilibrium of the game is not Pareto efficient, since isoprofit curves are not tangent. The outcome is below that of perfect competition and therefore is not socially optimal, but it is better than the monopoly outcome.

Extending the model to more than two firms, we can observe that the equilibrium of the game gets closer to the perfect competition outcome as the number of firms increases, decreasing market concentration.

### Illustration

Let the inverse demand function and the cost function be given by

$$P = 50 - 2Q \text{ and } C = 10 + 2q$$

respectively, where Q is total industry output and q is the firm's output.

First consider first the case of uniform-pricing monopoly, as a benchmark. Then in this case  $Q = q$  and the profit function is

$$\pi(Q) = (50 - 2Q)Q - 10 - 2Q = 48Q - 2Q^2 - 10.$$

Solving

$$\begin{aligned} \frac{d\pi}{dQ} = 0 \text{ we get } Q = 12, P = 26, \pi = 278, CS = \frac{12(50 - 26)}{2} = 144, TS \\ = 278 + 144 = 422. \end{aligned}$$

### MONOPOLY

Q	P	$\pi$	CS	TS
12	26	278	144	422

Now let us consider the case of two firms, or duopoly. Let  $q_1$  be the output of firm 1 and  $q_2$  the output of firm 2. Then  $Q = q_1 + q_2$  and the profit functions are:

$$\begin{aligned} \pi_1(q_1, q_2) &= q_1 [50 - 2(q_1 + q_2)] - 10 - 2q_1 \\ \pi_2(q_1, q_2) &= q_2 [50 - 2(q_1 + q_2)] - 10 - 2q_2 \end{aligned}$$

A Nash equilibrium is a pair of output levels  $(q_1^*, q_2^*)$  such that:

$$\pi_1(q_1^*, q_2^*) \geq \pi_1(q_1, q_2^*) \text{ for all } q_1 \geq 0$$

and

$$\pi_1(q_1^*, q_2^*) \geq \pi_1(q_1^*, q_2) \text{ for all } q_2 \geq 0$$

This means that, fixing  $q_2$  at the value  $q_2^*$  and considering  $\pi_1$  as a function of  $q_1$  alone, this function is maximized at  $q_1=q_1^*$ . But a necessary condition for this to be true is that

$$\frac{\partial \pi_1}{\partial q_1}(q_1^*, q_2^*) = 0.$$

Similarly, fixing  $q_1$  at the value  $q_1^*$  and considering  $\pi_2$  as a function of  $q_2$  alone, this function is maximized at  $q_2=q_2^*$ . But a necessary condition for this to be true is that

$$\frac{\partial \pi_2}{\partial q_2}(q_1^*, q_2^*) = 0.$$

Thus the Nash equilibrium is found by solving the following system of two equations in the two unknowns  $q_1$  and  $q_2$ :

$$\begin{cases} \frac{\partial \pi_1}{\partial q_1}(q_1^*, q_2^*) = 50 - 4q_1 - 2q_2 - 2 = 0 \\ \frac{\partial \pi_2}{\partial q_2}(q_1^*, q_2^*) = 50 - 2q_1 - 4q_2 - 2 = 0 \end{cases}$$

The solution is  $q_1^* = q_2^* = 8, Q = 16, P = 18, \pi_1 = \pi_2 = 118, CS = \frac{16(50-18)}{2} = 256, TS = 118 + 118 + 256 = 492$ .

Let us compare the two.

### MONOPOLY

Q	P	$\pi$	CS	TS
12	26	278	144	422

### DUOPOLY

$q_1$	$q_2$	Q	P	$\pi_1$	$\pi_2$	tot $\pi$	CS	TS
8	8	16	18	118	118	236	256	492

Thus competition leads to an increase not only in consumer surplus but in total surplus: the gain in consumer surplus ( $256 - 144 = 112$ ) exceeds the loss in total profits ( $278 - 236 = 42$ ).

In the above example we assumed that the two firms had the same cost function ( $C = 10 + 2q$ ). However, there is no reason why this should be true. The same reasoning applies to the case where the firms have different costs. Example: demand function as before ( $P = 50 - 2Q$ ) but now

$$\text{cost function of firm 1: } C_1 = 10 + 2q_1$$

$$\text{cost function of firm 2: } C_2 = 12 + 8q_2.$$

Then the profit functions are:

$$\pi_1(q_1, q_2) = q_1 [50 - 2(q_1 + q_2)] - 10 - 2q_1$$

$$\pi_2(q_1, q_2) = q_2 [50 - 2(q_1 + q_2)] - 12 - 8q_2$$

The Nash equilibrium is found by solving:

$$\begin{cases} \frac{\partial \pi_1}{\partial q_1}(q_1^*, q_2^*) = 50 - 4q_1 - 2q_2 - 2 = 0 \\ \frac{\partial \pi_2}{\partial q_2}(q_1^*, q_2^*) = 50 - 2q_1 - 4q_2 - 8 = 0 \end{cases}$$

The solution is  $q_1^* = 9$ ,  $q_2^* = 6$ ,  $Q = 15$ ,  $P = 20$ ,  $\pi_1 = 152$ ,  $\pi_2 = 60$ . Since firms have different costs, they choose different output levels: the low-cost firm (firm 1) produces more and makes higher profits than the high-cost firm (firm 2).

### SELF-ASSESSMENT EXERCISE

Consider the Cournot model when firm 1 has a cost advantage over firm 2, where  $c_1 < c_2$ . All other conditions remain the same so that firm profits are

$$\begin{aligned} \pi_1 &= aq_1 - bq_1^2 - bq_1q_2 - c_1q_1 \\ \pi_2 &= aq_2 - bq_2^2 - bq_1q_2 - c_2q_2 \end{aligned}$$

Obtain the Cournot equilibrium  $q_1^*$ ,  $q_2^*$ ,  $P^*$ ,  $\pi_1^*$  and  $\pi_2^*$ ,

**HINT:** solve the first-order conditions of profit maximization and solve them simultaneously for output, and plug these optimal values into the demand and profit functions to obtain the Cournot equilibrium.

## 2. The Bertrand Duopoly

Joseph Bertrand, operating around the same period as Cournot, was a French mathematician and economist. Bertrand became well known after publishing a number of reviews on math and economy-related articles written by professional peers and colleagues such as Leon Walras and Antoine Cournot.

Bertrand's critique of Cournot's model of duopolies is ultimately what led to the furthering of both oligopoly theory and game theory, most notably resulting in the formation of his own theory or model of duopolies, the Bertrand Model.

The primary difference between Cournot's model and Bertrand's model is that while Cournot believed production quantity would drive the competition between the two companies, Bertrand believed that the competition would always be driven by price.

Bertrand's duopoly theory identified that consumers, when given a choice between equal or similar goods and services, will opt for the company that gives the best price. This would start a price war, with both companies dropping prices, leading to an inevitable loss of profits.

The Cournot duopoly focus on firm competing through the quantity of output they produce. The Bertrand duopoly model examines price competition among firms that produce differentiated but highly substitutable products. Each firm's quantity demanded is a function of not only the price it charges but also the price charged by its rival. Coca-Cola and Pepsi are examples of Bertrand duopolists.

The quantity demanded for firm A and firm B is a function of both the price the firm establishes and the price established by their rival because the goods are highly substitutable. Thus, the firms have the following demand curves relating quantity demanded to its price and its rival's price

$$\text{Firm A: } q_A = 400 - 4P_A + 2P_B \dots \dots \dots 1$$

$$\text{Firm B: } q_B = 240 - 3P_B + 1.5P_A \dots \dots \dots 2$$

To simplify the analysis, assume that both firms have zero marginal cost for their products. Profit maximization then requires each firm to choose a price that maximizes its total revenue.

Derive the Bertrand reaction functions for each firm with the following steps:

1. Firm A's total revenue equals price times quantity, so

$$TR_A = P_A \times q_A = P_A(400 - 4P_A + 2P_B) = 400P_A - 4P_A^2 + 2P_AP_B$$

2. Taking the derivative of firm A's total revenue with respect to the price it charges yields

$$\frac{dTR_A}{dP_A} = 400 - 8P_A + 2P_B \dots \dots \dots 3$$

3. Setting the equation (3) equal to zero and solving it for  $P_A$  generates firm's A reaction function. Setting the derivative of total revenue equal to zero maximizes total revenue, which also maximizes profit given marginal cost equals zero.

$$\frac{dTR_A}{dP_A} = 400 - 8P_A + 2P_B = 0 \text{ or } 400 + 2P_B = 8P_A$$

$$P_A = 50 + 0.25P_B$$

4. Repeat these steps for firm B to derive its reaction function.

$$TR_B = P_B \times q_B = P_B(240 - 3P_B + 1.5P_A) = 240P_B - 3P_B^2 + 1.5P_AP_B$$

$$\frac{dTR_B}{dP_B} = 240 - 6P_B + 1.5P_A \dots \dots \dots 4$$

$$\frac{dTR_A}{dP_A} = 240 - 6P_B + 1.5P_A = 0 \text{ or } 240 + 1.5P_A = 6P_B$$

$$P_B = 40 + 0.25P_A$$

5. Substitute firm B's reaction function into firm A's reaction to determine  $P_A$ .

$$P_A = 50 + 0.25(40 + 0.25P_A) = 50 + 10 + 0.0625P_A = 60 + 0.0625P_A$$

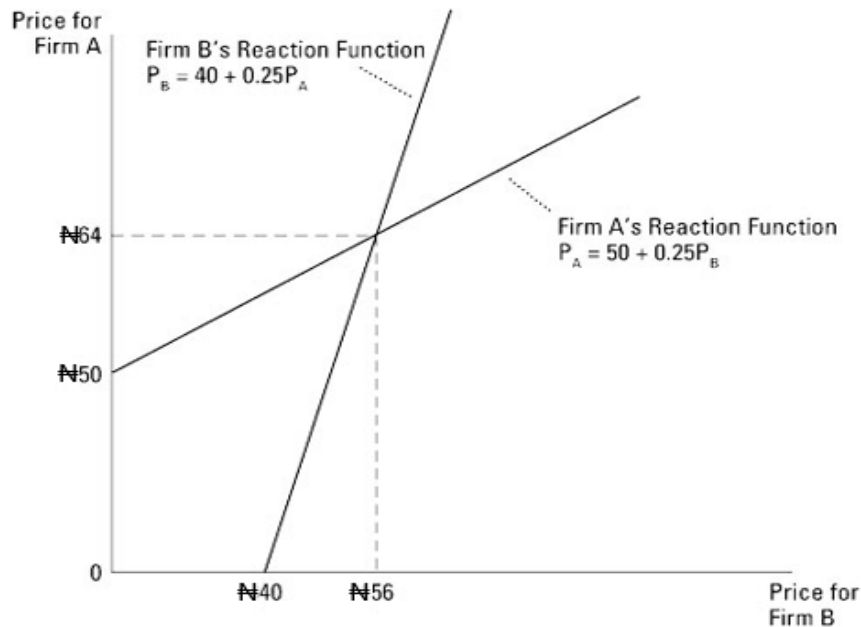
$$0.9375P_A = 60$$

$$P_A = 64$$

6. Substitute  $P_A$  equals 64 in firm B's reaction function to determine  $P_B$ .

$$P_B = 40 + 0.25(64) = 56$$

The Bertrand duopoly model indicates that firm A maximizes profit by charging ₦64, and firm B maximizes profit by charging ₦56. Note that both the horizontal and vertical axes on the illustration measure price and not quantity (as in the Cournot and Stackelberg models).



In the Bertrand model, firms compete with price. Therefore, reaction functions are expressed in terms of price, not quantities.

### SELF-ASSESSMENT EXERCISE

$$q_1 = \alpha - \beta p_1 + \delta p_2,$$

$$q_2 = \alpha - \beta p_2 + \delta p_1,$$

where  $\alpha = \frac{a(1-d)}{x}$ ,  $\beta = \frac{1}{x}$ ,  $\delta = \frac{d}{x}$ , and  $x = (1 - d^2)$ .

With this demand system, firm  $i$ 's profits are

$$\pi_i(p_i, p_j) = (p_i - c)q_i - F_i = (p_i - c)(\alpha - \beta p_i + \delta p_j) - F_i.$$

Derive each firm's first-order condition with respect to its own price for a given rival price and solve the system of equations. Thus, generates Bertrand equilibrium prices.

### 3.2.2 The Significance of a Duopoly

Duopolies are significant because they force each company to consider how its actions will affect its rival, meaning, how the rival firm will respond. It affects how each company operates, how it produces its goods, and how it advertises its services, and can ultimately change what and how goods and services are both offered and priced. When the two firms compete on price – in a Bertrand Duopoly – prices tend to dip to or below the cost of production, thereby wiping out any chance for profit.

For this reason, most duopolistic firms find it profitable and generally necessary to agree to form a sort of monopoly, setting prices that allow both firms to take one half of the market space and thus one half of the market's profit. However, this is a tricky tactic if done incorrectly, because the Sherman Act and other antitrust laws in the United States make collusive activity illegal.

Duopolies, when operating and competing based on production quantity instead of price, tend to function better, avoiding any potential for legal issues and enabling each

firm to share in the profits, reaching a price and operating homeostasis within their duopolistic market.

### **3.3 Oligopoly**

Many purchases that individuals make at the retail level are produced in markets that are neither perfectly competitive, monopolies, nor monopolistically competitive. Rather, they are oligopolies. Oligopoly arises when a small number of large firms have all or most of the sales in an industry. Examples of oligopoly abound and include the auto industry, cable television, and commercial air travel.

Oligopolistic firms are like cats in a bag. They can either scratch each other to pieces or cuddle up and get comfortable with one another. If oligopolists compete hard, they may end up acting very much like perfect competitors, driving down costs and leading to zero profits for all. If oligopolists collude with each other, they may effectively act like a monopoly and succeed in pushing up prices and earning consistently high levels of profit. Oligopolies are typically characterized by mutual interdependence where various decisions such as output, price, advertising, and so on, depend on the decisions of the other firm(s).

Analysing the choices of oligopolistic firms about pricing and quantity produced involves considering the pros and cons of competition versus collusion at a given point in time.

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 market 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.

#### **3.3.1 Strategic Interactions**

Each firm must consider both: (1) other firms' reactions to the firm's own decisions, and (2) own firm's reactions to the other firms' decisions. Thus, there is a continuous interplay between decisions and reactions to those decisions by all firms in the industry. Each oligopolist must take into account these strategic interactions when making decisions. Since all firms in an oligopoly have outcomes that depend on the other firms, these strategic interactions are the foundation of the study and understanding of oligopoly.

For example, each automobile firm's market share depends on the prices and quantities of all of the other firms in the industry. If Ford lowers prices relative to other car manufacturers, it will increase its market share at the expense of the other companies.

When making decisions that consider the possible reactions of other firms, firm managers usually assume that the manager of competing firms are rational and intelligent. The strategic interactions that ensue form the study of game theory. John Nash (1928-2015) was an American mathematician who was one of the pioneers in



game theory. Economists and mathematicians use the concept of a Nash Equilibrium to describe a common outcome in game theory that is frequently used in the study of oligopoly.

Nash Equilibrium = An outcome where there is no tendency to change based on each individual choosing a strategy given the strategy of rivals.

In the study of oligopoly, the Nash Equilibrium assumes that each firm makes rational profit-maximizing decisions while holding the behaviour of rival firms constant. This assumption is made to simplify oligopoly models, given the potential for enormous complexity of strategic interactions between firms. As an aside, this assumption is one of the interesting themes of the motion picture, “A Beautiful Mind,” starring Russell Crowe as John Nash. The concept of Nash Equilibrium is also the foundation of the models of oligopoly presented in the next three sections: the Cournot, Bertrand, and Stackelberg models of oligopoly.

### 3.3.2 Cournot Model

Augustin Cournot (1801-1877) was a French mathematician who developed the first model of oligopoly explored here. The Cournot model is a model of oligopoly in which firms produce a homogeneous good, and assumes that the competitor’s output is fixed when deciding how much to produce.

A numerical example of the Cournot model follows, where it is assumed that there are two identical firms (a duopoly), with output given by  $Q_i$  ( $i=1,2$ ). Therefore, total industry output is equal to:  $Q = Q_1 + Q_2$ . Market demand is a function of price and given by  $Q^d = Q^d(P)$ , thus the inverse demand function is  $P = P(Q^d)$ . Note that the price depends on the market output  $Q$ , which is the sum of both individual firm’s outputs. In this way, each firm’s output has an influence on the price and profits of both firms. This is the basis for strategic interaction in the Cournot model: if one firm increases output, it lowers the price facing both firms. The inverse demand function and cost function are given in the equation below:

$$\begin{aligned} P &= 40 - Q \\ C(Q_i) &= 7Q_i \\ i &= 1,2 \end{aligned}$$

Each firm chooses the optimal, profit-maximizing output level given the other firm’s output. This will result in a Nash Equilibrium, since each firm is holding the behaviour of the rival constant. Firm One maximizes profits as follows:

$$\begin{aligned} \max \pi_1 &= TR_1 - TC_1 \\ \max \pi_1 &= P(Q)Q_1 - C(Q_1) \text{ [price depends on total output } Q = Q_1 + Q_2\text{]} \\ \max \pi_1 &= [40 - Q]Q_1 - 7Q_1 \\ \max \pi_1 &= [40 - Q_1 - Q_2]Q_1 - 7Q_1 \\ \max \pi_1 &= 40Q_1 - Q_1^2 - Q_2Q_1 - 7Q_1 \\ \frac{\partial \pi_1}{\partial Q_1} &= 40 - 2Q_1 - Q_2 - 7 = 0 \\ 2Q_1 &= 33 - Q_2 \\ Q_1^* &= 16.5 - 0.5Q_2 \end{aligned}$$

This equation is called the “Reaction Function” of Firm One. This is as far as the mathematical solution can be simplified, and represents the Cournot solution for Firm One. It is a reaction function since it describes what Firm One’s reaction will be given the output level of Firm Two. This equation represents the strategic interactions between the two firms, as changes in Firm Two’s output level will result in changes in Firm One’s response. Firm One’s optimal output level depends on Firm Two’s behaviour and decision making. Oligopolists are interconnected in both behaviour and outcomes.

The two firms are assumed to be identical in this duopoly. Therefore, Firm Two’s reaction function will be symmetrical to the Firm One’s reaction function (check this by setting up and solving the profit-maximization equation for Firm Two):

$$Q_2^* = 16.5 - 0.5Q_1$$

The two reaction functions can be used to solve for the Cournot-Nash Equilibrium. There are two equations and two unknowns ( $Q_1$  and  $Q_2$ ), so a numerical solution is found through substitution of one equation into the other.

$$\begin{aligned} Q_1^* &= 16.5 - 0.5(16.5 - 0.5Q_1) \\ Q_1^* &= 16.5 - 8.25 + 0.25Q_1 \\ Q_1^* &= 8.25 + 0.25Q_1 \\ Q_1^* &= 8.25 + 0.25Q_1 \\ 0.75Q_1^* &= 8.25 \\ Q_1^* &= 8.25 \\ Q_1^* &= 11 \end{aligned}$$

Due to symmetry from the assumption of identical firms:

$$Q_i = 11; i = 1, 2; Q = 22; P = 18$$

Profits for each firm are:

$$\pi_i = P(Q)Q_i - C(Q_i) = 18(11) - 7(11) = (18 - 7)11 = 11(11) = \text{€}121$$

This is the Cournot-Nash solution for oligopoly, found by each firm assuming that the other firm holds its output level constant. The Cournot model can be easily extended to more than two firms, but the math does get increasingly complex as more firms are added. Economists utilize the Cournot model because it is based on intuitive and realistic assumptions, and the Cournot solution is intermediary between the outcomes of the two extreme market structures of perfect competition and monopoly.

This can be seen by solving the numerical example for competition, Cournot, and monopoly models, and comparing the solutions for each market structure.

In a competitive industry, free entry results in price equal to marginal cost. In the case of the numerical example,  $P = 7$ . When this competitive price is substituted into the inverse demand equation,  $7 = 40 - Q$ , or  $Q_c = 33$ . Profits are found by solving  $(P - MC)Q$ , or  $\pi_c = (7 - 7)Q = 0$ . The competitive solution is given in the equation below:

$$P_c = 7; \quad Q_c = 33; \quad \pi_c = 0$$

The monopoly solution is found by maximizing profits as a single firm.

$$\begin{aligned} \max \pi_m &= TR_m - TC_m \\ \max \pi_m &= P(Q_m)Q_m - C(Q_m) \text{ [price depends on total output } Q = Q_1 + Q_2] \\ \max \pi_m &= [40 - Q_m]Q_m - 7Q_m \\ \max \pi_m &= 40Q_m - Q_m^2 - 7Q_m \end{aligned}$$

$$\frac{\partial \pi_m}{\partial Q_m} = 40 - 2Q_m - 7 = 0$$

$$2Q_m = 33$$

$$Q_m^* = 16.5$$

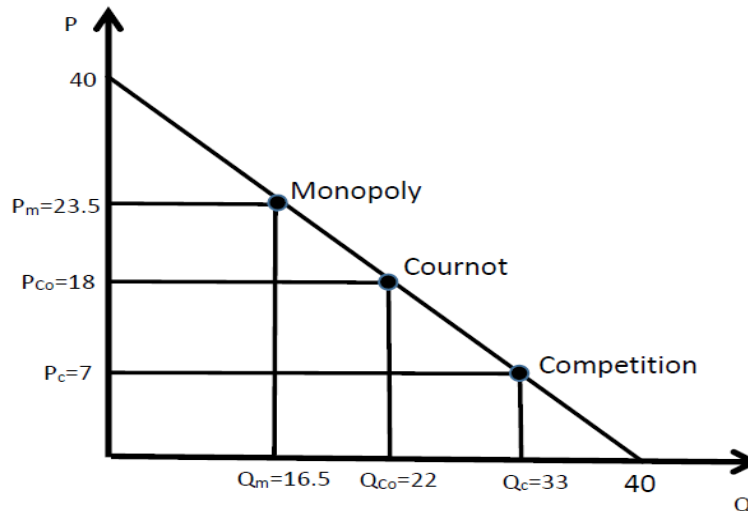
$$P_m = 40 - 16.5 = 23.5$$

$$\pi_m = (P_m - MC_m)Q_m = (23.5 - 7)16.5 = 16.5(16.5) = \text{₦}272.25$$

The monopoly solution is given in the equation below;

$$P_m = 23.5; \quad Q_m = 16.5; \quad \pi_m = \text{₦}272.5$$

The competitive, Cournot, and monopoly solutions can be compared on the same graph for the numerical example in the figure below;



**Figure 3.12: Comparison of Perfect Competition, Cournot, and Monopoly Solutions**

The Cournot price and quantity are between perfect competition and monopoly, which is an expected result, since the number of firms in an oligopoly lies between the two market structures extremes.

### 3.3.3 The Bertrand Model

This model addresses the case of not so many numbers 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.

The Bertrand model is a model of oligopoly in which firms produce a homogeneous good, and each firm takes the price of competitors fixed when deciding what price to charge.

Assume two firms in an oligopoly (a duopoly), where the two firms choose the price of their good simultaneously at the beginning of each period. Consumers purchase from the lowest price firm, since the products are homogeneous (perfect substitutes). If the two firms charge the same price, one-half of the consumers buy from each firm. Let the demand equation be given by  $Q^d = Q^d(P)$ . The Bertrand model follows these three statements:

- (1) If  $P_1 < P_2$ , then Firm One sells  $Q^d$  and Firm Two sells 0.
- (2) If  $P_1 > P_2$ , then Firm One sells 0 and Firm Two sells  $Q^d$ .
- (3) If  $P_1 = P_2$ , then Firm One sells  $0.5Q^d$  and Firm Two sells  $0.5Q^d$ .

A numerical example demonstrates the outcome of the Bertrand model, which is a Nash Equilibrium. Assume two firms sell a homogeneous product, and compete by choosing prices simultaneously while holding the other firm's price constant. Let the demand function be given by  $Q^d = 50 - P$  and the costs are summarized by  $MC_1 = MC_2 = 5$ .

- (1) Firm One sets  $P_1 = 20$ , and Firm Two sets  $P_2 = 15$ . Firm Two has the lower price, so all customers purchase the good from Firm Two.

$$Q_1 = 0, Q_2 = 35. \pi_1 = 0, \pi_2 = (15 - 5)35 = \text{₦}350.$$

After period one, Firm One has a strong incentive to lower the price below  $P_2$ . The Bertrand assumption is that both firms will choose a price, holding the other firm's price constant. Thus, Firm One undercuts  $P_2$  slightly, assuming that Firm Two will maintain its price at  $P_2 = 15$ . Firm Two will keep the same price, assuming that Firm One will maintain  $P_1 = 20$ .

- (2) Firm One sets  $P_1 = 14$ , and Firm Two sets  $P_2 = 15$ . Firm One has the lower price, so all customers purchase the good from Firm One.

$$Q_1 = 36, Q_2 = 0. \pi_1 = (14 - 5)36 = \text{₦}324, \pi_2 = 0.$$

After period two, Firm Two has a strong incentive to lower price below  $P_1$ . This process of undercutting the other firm's price will continue and the "price war" will result in the price being driven down to marginal cost. In equilibrium, both firms lower their price until price is equal to marginal cost:  $P_1 = P_2 = MC_1 = MC_2$ . The price cannot go lower than this, or the firms would go out of business due to negative economic profits. To restate the Bertrand model, each firm selects a price, given the other firm's price.

The Bertrand results are given in the Equation below:

$$\begin{aligned} P_1 &= P_2 = MC_1 = MC_2 \\ Q_1 &= Q_2 = 0.5Q^d \\ \pi_1 &= \pi_2 = 0 \text{ in the SR and LR.} \end{aligned}$$

The Bertrand model of oligopoly suggests that oligopolies have the competitive solution, due to competing over price. There are many oligopolies that behave this way, such as gasoline stations at a given location. Other oligopolies may behave more like Cournot oligopolists, with an outcome somewhere in between perfect competition and monopoly.

### 3.3.4 Stackelberg Model

Heinrich Freiherr von Stackelberg (1905-1946) was a German economist who contributed to game theory and the study of market structures with a model of firm leadership, or the Stackelberg model of oligopoly. This model assumes that there are two firms in the industry, but they are asymmetrical: there is a "leader" and a "follower." Stackelberg used this model of oligopoly to determine if there was an advantage to going first, or a "first-mover advantage."

A numerical example is used to explore the Stackelberg model. Assume two firms, where Firm One is the leader and produces  $Q_1$  units of a homogeneous good. Firm Two

is the follower, and produces  $Q_2$  units of the good. The inverse demand function is given by  $P = 100 - Q$ , where  $Q = Q_1 + Q_2$ . The costs of production are given by the cost function:  $C(Q) = 10Q$ .

This model is solved recursively, or backwards. Mathematically, the problem must be solved this way to find a solution. Intuitively, each firm will hold the other firm's output constant, similar to Cournot, but the leader must know the follower's best strategy in order to move first. Thus, Firm One solves Firm Two's profit maximization problem in order to know what output it will produce, or Firm Two's reaction function. Once the reaction function of the follower (Firm Two) is known, then the leader (Firm One) maximizes profits by substitution of Firm Two's reaction function into Firm One's profit maximization equation.

Firm One starts by solving for Firm Two's reaction function:

$$\begin{aligned} \max \pi_2 &= TR_2 - TC_2 \\ \max \pi_2 &= P(Q)Q_2 - C(Q_2) \text{ [price depends on total output } Q = Q_1 + Q_2\text{]} \\ \max \pi_2 &= [100 - Q]Q_2 - 10Q_2 \\ \max \pi_2 &= [100 - Q_1 - Q_2]Q_2 - 10Q_2 \\ \max \pi_2 &= 100Q_2 - Q_1Q_2 - Q_2^2 - 10Q_2 \\ \frac{\partial \pi_2}{\partial Q_2} &= 100 - Q_1 - 2Q_2 - 10 = 0 \\ 2Q_2 &= 90 - Q_1 \\ Q_2^* &= 45 - 0.5Q_1 \end{aligned}$$

Next, Firm One, the leader, maximizes profits holding the follower's output constant using the reaction function:

$$\begin{aligned} \max \pi_1 &= TR_1 - TC_1 \\ \max \pi_1 &= P(Q)Q_1 - C(Q_1) \text{ [price depends on total output } Q = Q_1 + Q_2\text{]} \\ \max \pi_1 &= [100 - Q]Q_1 - 10Q_1 \\ \max \pi_1 &= [100 - Q_1 - Q_2]Q_1 - 10Q_1 \\ \max \pi_1 &= [100 - Q_1 - (45 - 0.5Q_1)]Q_1 - 10Q_1 \text{ [substitution of One's reaction function]} \\ \max \pi_1 &= [100 - Q_1 - 45 + 0.5Q_1]Q_1 - 10Q_1 \\ \max \pi_1 &= [55 - 0.5Q_1]Q_1 - 10Q_1 \\ \max \pi_1 &= 55Q_1 - 0.5Q_1^2 - 10Q_1 \\ \frac{\partial \pi_1}{\partial Q_1} &= 55 - Q_1 - 10 = 0 \\ Q_1^* &= 45 \end{aligned}$$

This can be substituted back into Firm Two's reaction function to solve for  $Q_2^*$ .

$$\begin{aligned} Q_2^* &= 45 - 0.5Q_1 = 45 - 0.5(45) = 45 - 22.5 = 22.5 \\ Q &= Q_1 + Q_2 = 45 + 22.5 = 67.5 \\ P &= 100 - Q = 100 - 67.5 = 32.5 \\ \pi_1 &= (32.5 - 10)45 = 22.5(45) = \text{€}1,012.5 \\ \pi_2 &= (22.5 - 10)22.5 = 12.5(22.5) = \text{€}506.25 \end{aligned}$$

#### 4.0 CONCLUSION

Cournot's duopoly: In this model, the firms simultaneously choose quantities. Bertrand's oligopoly: In this model, the firms simultaneously choose prices.

We covered three models of oligopoly: Cournot, Bertrand, and Stackelberg. These three models are alternative representations of oligopolistic behaviour. The Bertrand model is relatively easy to identify in the real world, since it results in a price war and competitive prices. It may be more difficult to identify which of the quantity models to use to analyze a real-world industry: Cournot or Stackelberg?

The model that is most appropriate depends on the industry under investigation. The Cournot model may be most appropriate for an industry with similar firms, with no market advantages or leadership. While the Stackelberg model may be most appropriate for an industry dominated by relatively large firms.

Oligopoly has many different possible outcomes, and several economic models to better understand the diversity of industries. Notice that if the firms in an oligopoly colluded, or acted as a single firm, they could achieve the monopoly outcome. If firms banded together to make united decisions, the firms could set the price or quantity as a monopolist would. This is illegal in many nations, including the United States, since the outcome is anti-competitive, and consumers would have to pay monopoly prices under collusion.

If firms were able to collude, they could divide the market into shares and jointly produce the monopoly quantity by restricting output. This would result in the monopoly price, and the firms would earn monopoly profits. However, under such circumstances, there is always an incentive to “cheat” on the agreement by selling more output. If the other firms in the industry are restricting output, a firm can increase profits by increasing output, at the expense of the other firms in the collusive agreement. We will discuss this possibility in the next section.

## **5.0 SUMMARY**

The production functions describe 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 was discussed; the effect of technological progress on production was also discussed.

An oligopoly is characterized by a market with a few firms that compete in a strategic setting because a firm’s profits and best course of action depend on the actions of its competitors.

A duopoly is a special case of oligopoly where there are two firms in the market.

In the Cournot model, each firm chooses a level of output that maximizes its profits, given the output of its competitors. The Cournot equilibrium is a Nash equilibrium where each firm correctly assumes that its competitors behave optimally. As the number of firms in a market changes from 1 to many, the Cournot equilibrium changes from monopoly to the perfectly competitive equilibrium. In the Bertrand model with homogeneous goods, each firm chooses its price to maximize profits, given the price of its competitors. The Bertrand equilibrium is a Nash equilibrium where the price equals the perfectly competitive price as long as there are two or more firms in the market. This occurs because each firm will undercut the price of its rivals until the competitive price is reached. This outcome is called the Bertrand Paradox.

The most important characteristic of oligopoly is that firm decisions are based on strategic interactions. Each firm’s behavior is strategic, and strategy depends on the

other firms' strategies. Therefore, oligopolists are locked into a relationship with rivals that differs markedly from perfect competition and monopoly.

## 6.0 TUTOR-MARKED ASSIGNMENT

1. Consider a market with two firms (1 and 2) that face a linear inverse demand function  $p = a - bQ$  and a total cost function  $TC = cq_i, c > 0$ . Find the Cournot equilibrium output for each firm. How will your answer change if  $TC = c_2q_i$ ?
2. Consider a market with two firms (1 and 2) that face a linear demand function  $Q = 24 - p$  and a total cost function  $TC = cq_i, c > 0$ . Find the Bertrand equilibrium price. How will your answer change if  $c_1 = 10$  and  $c_2 = 12$ ?
3. Assume a duopoly market where firms can choose to compete in output or in price. Provide a simple example where it is optimal for both firms to compete in price instead of output.
4. Suppose that the (inverse) demand curve for Viagra is given by  $P = 200 - 2Q$  Where  $Q$  is total industry output. The market is occupied by two firms, each with constant marginal costs equal to ₹8.
  - a) Calculate the equilibrium price and quantity assuming the two firms compete in quantities.
  - b) How would your answer to (a) change if one of the firm's costs rose to ₹10?
  - c) Repeat parts (a) and (b) assuming the competition is in prices rather than quantities.
5. Suppose that the (inverse) market demand for fax paper is given by  $P = 400 - 2Q$ . Where  $Q$  is total industry output. There are two firms that produce fax paper. Each firm has a constant marginal cost of production equal to ₹40 and they are competing in quantities. That is, they each choose production levels simultaneously.
  - a) Calculate the best response function for each firm (i.e. each firm's profit maximizing choice of quantity given the other firm's production levels)
  - b) Calculate the profit maximizing price/quantity for a monopolist facing the same demand curve (and with the same production costs). How does your answer compare to (a)?

## 7.0 REFERENCES/FURTHER READING

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

Jhingan M.L (2009): *Microeconomics Theory*, Vrinda Publications (P) Limited, Delhi

Koutsoyiannis, A. (2001). *Modern Microeconomics* (2nd ed.). Hampshire: Macmillan Press Ltd.

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

Robert J. Graham, “How to Compete for Customers: The Bertrand Model of Duopolies in Managerial Economics”. Retrieved on the 22/04/2020 from <https://www.dummies.com/education/economics/how-to-compete-for-customers-the-bertrand-model-of-duopolies-in-managerial-economics/>

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



## **UNIT 2: BILATERAL MONOPOLY AND MONOPSONY**

### **CONTENTS**

- 1.0 Introduction
- 2.0 Objectives
- 3.0 Main Content
  - 3.1 Price Discrimination
  - 3.2 Bilateral Monopoly
  - 3.3 Monopsony
- 4.0 Conclusion
- 5.0 Summary
- 6.0 Tutor-Marked Assignment
- 7.0 References/Further Reading

### **1.0 INTRODUCTION**

#### **2.0 OBJECTIVES**

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

- ✓ Define the terms bilateral monopoly and monopsony
- ✓ How firms determine wages and employment when a specific labour market combines a union and a monopsony
- ✓ Analyse monopsony, and bilateral monopoly in the Nigeria context.

### **3.0 MAIN CONTENT**

#### **3.1 Bilateral Monopoly**

A bilateral monopoly is a market that is characterized by one firm or individual, a monopolist, on the supply side and one firm or individual, a monopsonist, on the demand side. The input markets of the monopolist and the output market of the monopsonist can be of any form. The essential ingredient is the single seller-single buyer situation. Because a buyer and a seller of a product, perforce, do business with each other, they are clearly able to make legally binding agreements. This contrasts with firms in the same industry, which do not sell to one another, and which are often precluded by anti-collusion laws from making legally enforceable contracts. Of course, it is also possible to view bilateral monopoly non-cooperatively.

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, 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 monopolise 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-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.

The one supplier will tend to act as a monopoly power and look to charge high prices to the one buyer. The lone buyer will look towards paying a price that is as low as possible. Since both parties have conflicting goals, the two sides must negotiate based on the relative bargaining power of each, with a final price settling in between the two sides' points of maximum profit.

This climate can exist whenever there is a small contained market, which limits the number of players, or when there are multiple players but the costs to switch buyers or sellers is prohibitively expensive.

### **3.1.1 How a Bilateral Monopoly works**

Bilateral monopoly requires the seller and the buyer, who have diametrically opposite interests, to achieve a balance of their interests. The buyer seeks to buy cheap, and the seller tries to sell expensive. The key to a successful business for both is reaching a balance of interests reflected in a “win-win” model. At the same time, both the seller and the buyer are well aware of who they are dealing with.

### **3.1.2 Disadvantages of Bilateral Monopoly**

Problems arise when neither party can determine the conditions of sale, nor the negotiation goes beyond what is permissible. For example, instead of fair negotiation and exchanging draft contracts, the buyer and seller abuse their rights: they stop shipping goods, impose unprofitable and discriminatory conditions, send false information to each other, etc. This creates uncertainty and threatens the entire market.

A common type of a bilateral monopoly occurs in a situation where there is a single large employer in a factory town, where its demand for labour is the only significant one in the city, and the labour supply is managed by a well-organized and strong trade union.

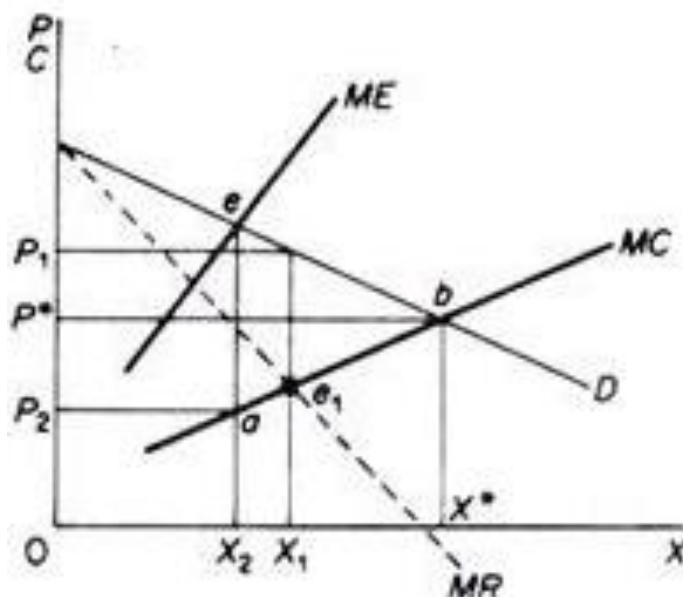
In such situations, the employer has no supply function that adequately describes the relationship between supply volume and product price. Therefore, the company must arbitrarily select a point on the market demand curve that maximizes his profit. The problem is that businesses in this situation are the only buyers of a monopolized product.

Consequently, its demand function for production resources is eliminated. Thus, to maximize his profit, the business must also choose a point on the seller's supply curve.

### **Illustration**

To illustrate a situation of bilateral monopoly assume that all railway equipment is produced by a single firm and is bought by a single buyer, British Rail. Both firms are assumed to aim at the maximization of their profit. The equilibrium of the producer-monopolist is defined by the intersection of his marginal revenue and marginal cost

curves (point  $e_1$  in figure 3.1.2). He would maximize his profit if he were to produce  $X_1$  quantity of equipment and sell it at the price  $P_1$ .



**Figure 3.13**

However, the producer cannot attain the above profit-maximizing position, because he does not sell in a market with many buyers, each of whom would be unable to affect the price by his purchases. The producer-monopolist is selling to a single buyer who can obviously affect the market price by his purchasing decisions.

The buyer is aware of his power, and, being a profit maximiser, he would like to impose his own price terms to the producer. What are the monopsonist's price terms? Clearly the MC curve of the producer represents the supply curve to the buyer: the upward slope of this curve shows that as the monopsonist increases his purchases the price he will have to pay rises. The MC (= S) curve is determined by conditions outside the control of the buyer, and it shows the quantity that the monopolist-seller is willing to supply at various prices.

The increase in the expenditure of the buyer (his marginal outlay or marginal expenditure) caused by the increases in his purchases is shown by the curve ME in figure 3.1.2. In other words, curve ME is the marginal cost of equipment for the monopsonist-buyer (it is a marginal-outlay curve to the total-supply curve MC, with which the buyer is faced). The equipment is an input for the buyer.

Thus in order to maximise his profit he would like to purchase additional units of X until his marginal outlay is equal to his price, as determined by the demand curve DD. The equilibrium of the monopsonist is shown by point e in figure 3.1.2 he would like to purchase  $X_2$  units of equipment at a price  $P_2$ , determined by point a on the supply curve MC(= S).

However, the monopsonist does not buy from a lot of small firms which would be price-takers (that is, who would accept the price imposed by the single buyer), but from the monopolist, who wants to charge price  $P_1$ . Given that the buyer wants to pay  $P_2$  while the seller wants to charge  $P_1$ , there is indeterminacy in the market. The two firms will sooner or later start negotiations and will eventually reach an agreement about price,

which will be settled somewhere in the range between  $P_1$  and  $P_2$ , ( $P_2 \leq P \leq P_1$ ), depending on the bargaining skill and power of the firms.

It should be obvious that a bilateral monopoly is rare for commodity markets, but is quite common in labour markets, where workers are organised in a union and confront a single employer (for example, the miners' unions and the Coal Board) or firms organised in a trade association.

If a bilateral monopoly emerges in a commodity market the buyer may attempt to buy out the seller-monopolist, thus attaining vertical integration of his production. The consequences of such take-over are interesting. The supply curve  $MC(=S)$  becomes the marginal-cost curve of the monopsonist, and hence his equilibrium will be defined by point b in figure 3.1.2 (where the 'new' marginal-cost curve intersects the price- demand curve  $DD$ ) output will increase to the level  $X$  and the marginal cost will be  $P_1$  lower than the price  $P$ , that the ex-monopolist would like to charge.

The result of the vertical integration in these conditions is an increase in the production of the input, which will lead to an increase in the final product of the ex-monopsonist and a reduction in his price, given that he is faced by a downward-sloping market-demand curve. The examination of the welfare implications of such a situation is beyond the scope of this elementary analysis.

### **SELF-ASSESSMENT EXERCISE**

What do you understand by the term bilateral monopoly?

### **3.2 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.

The three key characteristics of monopsony are:

- i. A single firm buying all output in a market
- ii. No alternative buyers
- iii. Restrictions on entry into the industry.

**i. 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.

**ii. 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.

- iii. **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:
- Government license or franchise
  - Resource ownership
  - Patents and copyrights
  - High start-up cost
  - Decreasing average total cost.

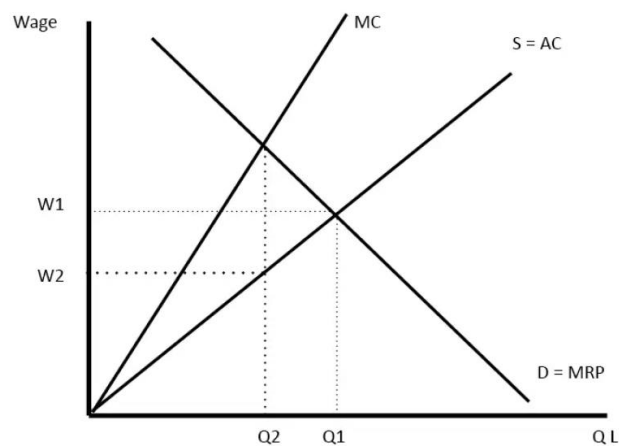
A monopsony occurs when a firm has market power in employing factors of production (e.g. labour). A monopsony means there is one buyer and many sellers. It often refers to a monopsony employer – who has market power in hiring workers. This is a similar concept to monopoly where there is one seller and many buyers.

### 3.2.1 Monopsony in Labour Markets

An example of a monopsony occurs when there is one major employer and many workers seeking to gain employment. If there is only one main employer of labour, then they have market power in setting wages and choosing how many workers to employ. Examples of monopsony in labour markets Coal mine owner in town where coal mining is the primary source of employment.

The government in the employment of civil servants, nurses, police and army officers.

**Diagram of monopsony**



**Figure 3.14: Monopsony**

In a competitive labour market, the equilibrium will be where  $D=S$  at  $Q1, W1$ . However, a monopsony can pay lower wages ( $W2$ ) and employ fewer workers ( $Q2$ )

### 3.2.2 Profit Maximisation for a Monopsony

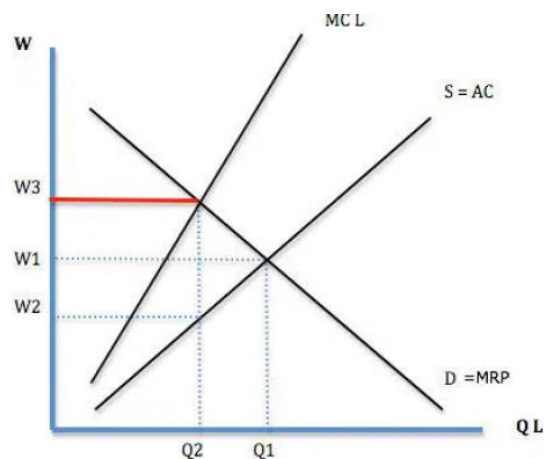
The marginal cost of employing one more worker will be higher than the average cost – because to employ one extra worker the firm has to increase the wages of all workers.

To maximise the level of profit, the firm employs  $Q2$  of workers where the marginal cost of labour equals the marginal revenue product  $MRP = D$

In a competitive labour market, the firm would be a wage taker. If they tried to pay only  $W2$ , workers would go to other firms willing to pay a higher wage.

### 3.2.3 Minimum Wage in a Monopsony

In a monopsony, a minimum wage can increase wages without causing unemployment.



A monopsony pays a wage of  $W_2$  and employs  $Q_2$ . If a minimum wage was placed equal to  $W_1$ , it would increase employment to  $Q_1$ . A minimum wage of  $W_3$  would keep employment at  $Q_2$ .

Even if a firm is not a pure monopsony, it may have a degree of monopsony power, due to geographical and occupational immobilities, which make it difficult for workers to switch jobs and find alternative employment.

For example, there are several employers who might employ supermarket checkout workers. However, in practice, it is difficult for workers to switch jobs to take advantage of slightly higher wages in other supermarkets. There is a lack of information and barriers to moving jobs. Therefore, although there are several buyers of labour, in practice the big supermarkets have a degree of monopsony power in employing workers.

### Illustration

Let's suppose we have a firm which is a monopsonist. To simplify matters, let's assume there is only factor of production: labour. In the competitive case, we have:

$$PQ(L) - wL$$

This implicitly assumes that the supply of labour to the firm is independent of the wage. In other words, labour supply is completely elastic. Suppose it was not perfectly elastic. Then you would have:

$$PQ(L(w)) - wL(w)$$

Alternatively we could write everything in terms of the inverse labour supply curve:

$$PQ(L) - w(L)L$$

Taking this second formulation, we get:

$$P \frac{dQ}{dL} - \frac{dw}{dL}L - w(L)$$

$$P \frac{dQ}{dL} = w(L) + \frac{dw}{dL}L$$

Now remember that  $\frac{dw}{dL} > 0$  (or at least this is usually what is assumed - people need a higher wage rate to induce them to work longer hours)

Then the marginal revenue product (the additional amount of revenue from an additional worker) is greater than the wage:

$$P \frac{dQ}{dL} > w(L)$$

This means that labor allocation is lower than in the absence of monopoly and thus output is also lower.

If the firm could hire an additional worker without it causing an increase in the wage paid to all the other workers, the firm would do so. Alternatively, the value created by an additional worker is greater than the wage the worker would receive and yet the firm is not willing to hire the worker. It would be societally worthwhile to increase production. However, the firm must also transfer some of its surplus to workers by raising wages when it increases output and so it chooses not to.

We can show this on a graph with labour supply, labour demand and the marginal expenditure curve:

$$\text{Labor Demand : } P \frac{dQ}{dL}$$

$$\text{Labor Supply : } L(w)$$

$$\text{Marginal Expenditure : } w(L) + \frac{dw}{dL}$$

### **Illustration 2**

Suppose that the price of a good is  $P = 10$  and the production function is constant returns to scale  $Q(L) = 2L$ : Lastly, assume that labour supply to the firm is given by:  $L = 50w$ :

Note that a higher wage increases labour supply. Solving for the wage as a function of labour, we get:

$$w = \frac{L}{50}$$

We can now write the profit function:

$$\pi(L) = PQ(L) - w(L)L = 10 * 2L - \frac{L}{50}L$$

Taking first order conditions from the profit function, we get:

$$\begin{aligned} \frac{d\pi}{dL} &= 20 - \frac{L}{25} = 0 \\ \Rightarrow L^* &= 500 \end{aligned}$$

From the labour supply equation, we can now solve for the wage:

$$w^* = \frac{L}{50} = 10$$

Now suppose that the firm was forced to pay 20 per hour (marginal revenue product of a worker or the marginal value of a worker):

The firm would then earn zero profits no matter what so it would be indifferent as to how many workers to hire:

$$\pi = 20L - 20L = 0$$

The amount of labour supplied would be determined by the labour supply curve:

$$L = 50w = 50 * 20 = 1000$$

Notice that the monopsonist keeps the wage below marginal revenue product. In a competitive environment, workers would get 20 but the monopsonist keeps the wage at 10: Moreover, in order to keep the wage low, the monopsonist has to restrict employment down to 500 instead of 1000: This also restricts output.

### **3.2.4 Problems of Monopsony in Labour Markets**

Monopsony can lead to lower wages for workers. This increases inequality in society. Workers are paid less than their marginal revenue product.

Firms with monopsony power often have a degree of monopoly selling power. This enables them to make high profits at the expense of consumers and workers.

Firms with monopsony power may also care less about working conditions because workers don't have many alternatives to the main firm. Monopsony in product markets. In several industries, there is one buyer and several sellers.

Supermarkets have monopsony power in buying food from farmers. If farmers don't sell to the big supermarkets, there are few alternatives. This has led to farmer protests about the price of milk.

Amazon.com is one of the biggest purchasers of books. If publishers don't sell to Amazon at a discounted price, they will miss out on selling to the biggest distributor of books.

### **SELF-ASSESSMENT EXERCISE**

What are the characteristics of monopsony?

## **4.0 CONCLUSION**

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

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

Bilateral monopoly is a market for a single buyer and seller or the combination of a monopoly market and a monopsony market.

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.



A bilateral monopoly is a labour market with a union on the supply side and a monopsony on the demand side. Since both sides have monopoly power, the equilibrium level of employment will be lower than that for a competitive labour market, but the equilibrium wage could be higher or lower depending on which side negotiates better. The union favours a higher wage, while the monopsony favours a lower wage, but the outcome is indeterminate in the model.

## 5.0 SUMMARY

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 maximising monopoly tends to charge a higher price, and a market dominated by a profit maximising monopsony tends to pay a lower price, thus when combined into a bilateral monopoly, the buyer and the seller both cannot maximise 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.

## 6.0 TUTOR-MARKED ASSIGNMENT

1. Analyse how the single buyers and sellers interact in a bilateral monopoly market.
2. Give an example of a monopsony market in Nigeria.
3. Suppose that in the labour market for soccer players we have the following equations: Supply Curve:  $W = \text{₦}4,050,000 + 2000L$ ; Demand Curve:  $MRP = \text{₦}6,000,000 - 1000L$  Marginal Cost Curve:  $MC = \text{₦}4,050,000 + 4000L$ 
  - (a) If the labour market for soccer players was competitive, determine the equilibrium wage (W) and employment level.
  - (b) If the labour market for soccer players was controlled by a monopsony determine the wage (W) and employment level (L).
4. The employment of teaching assistants (TAs) National Open University of Nigeria can be characterized as a monopsony. Suppose the demand for TAs is  $W = 30,000 - 125n$ , where W is the wage (as an annual salary), and n is the number of TAs hired. The supply of TAs is given by  $W = 1,000 + 75n$ .
  - a. If the university takes advantage of its monopsonist position, how many TAs will it hire? What wage will it pay?
  - b. If, instead, the university faced an infinite supply of TAs at the annual wage level of ₦10,000, how many TAs would it hire?

## 7.0 REFERENCES/FURTHER READING

- Case, et al. (2014). *Principles of Economics*. (11th ed.). Upper Saddle River, N J: 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 Extensions for Less-Developed Economies*.
- Jhingan M.L (2009): *Microeconomics Theory*, Vrinda Publications (P) Limited, Delhi

- Koutsoyiannis, A. (2001). *Modern Microeconomics* (2nd ed.). Hampshire: Macmillan Press Ltd.
- Robert, S. P. & Daniel, L. R. (2009). *Microeconomics*. (7th ed.). New York: Pearson Education International.

## **UNIT 3: THEORIES OF DETERMINATION OF WAGES, RENT, INTEREST AND PROFIT**

### **CONTENTS**

- 1.0 Introduction
- 2.0 Objectives
- 3.0 Main Content
  - 3.1 Labour market
  - 3.2 Supply and Demand of Labour
  - 3.3 The Demand for Labour and Wage Determination
    - 3.3.1 Human Capital Theory
  - 3.4 Determination of Rent
    - 3.4.1 Some Complications in the Rent of Land
  - 3.5 Determination of Interest
    - 3.5.1 Investment Production Process
    - 3.5.2 The Demand and Supply of Funds
  - 3.6 Determination of profit
    - 3.6.1 Profit Taxation
- 4.0 Conclusion
- 5.0 Summary
- 6.0 Tutor-Marked Assignment
- 7.0 References/Further Reading

### **1.0 INTRODUCTION**

In economics, factor payments are the rewards accruable to the various factors of production. These are the rewards for their participation in the process of production. These rewards are categorized into rent, wage, interest, and profit as the rewards for land, labour, capital and entrepreneurship respectively.

An economy is dependent on the production of goods and services hence factors of production are required for the production of goods and services. Land is the primary factor of production. Labour is the specific factor of production and payment is made in the form of wage. Capital is regarded as secondary factor of production as it can be manipulated by economic activities. Payment received would be in the form of interest. Later entrepreneurship was added as the fourth factor of production. It earns profit. Factors of production are owned by households and they supply these factors of production to firms and in return earn wages, interest, rent and profit. Households buy goods and services with this money.

### **2.0 OBJECTIVES**

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

- ✓ Understand how wage is determined
- ✓ Explain rent and other complications associated with it.
- ✓ Determine interest rate and profit

### 3.0 MAIN CONTENT

#### 3.1 Labour market

A firm will hire only the optimal amount of labour that will maximize its profit. This optimal quantity of labour depends on the marginal product of labour (MPL). The marginal product of labour is defined as the extra unit of output that the firm produces from hiring one extra unit of labour:

$$MPL = F(K, L + 1) - F(K, L)$$

The above equation states that the marginal product of labour is the difference of output produced from one extra unit of labour and the output produced from original quantity of labour. This production function implies diminishing marginal product which means that the marginal product of labour decreases as the amount of labour increases. As the amount of labour increases the production function becomes flatter i.e. diminishing marginal product.

#### 3.2 Supply and Demand of Labour

The demand for labour is dependent on the theory of marginal product of labour, which means that the firm compares the extra revenue earned from increased production that results from the added labour which ultimately leads to the higher spending on wages. Let us understand with the following example Let  $w$  denote wages rate that the labour receives,  $l^s$  denote the hours of labor services that household supplies to labour market. Thus the households receive  $wl^s$  amount of income from providing labour services. Now let  $l^d$  denote the labour services that the firm demands from labour market. Thus the firm pays  $wl^d$  amount of wage for providing labour services. Quantity produced in the economy will be given by  $y^s = f(l)^d$ . Thus the firm's profit will be given by gross revenue less wage payments.

$$profit = Py^s - wl^d = Pf(l)^d - wl^d$$

Now if we increase the labour input, it will have two effects. First, the output will increase by the marginal product of labour as extra hour of work is provided. Therefore, the gross sales revenue will be given by  $P.MPL$  second the wage rate for the firm increases by  $w$ . It is clear that in order to maximize profit, the firm will expand the employment level up to the level where the marginal product just equals wage rate.  $P.MPL = w$ . Now if we divide the above equation with the general price level, then we obtain  $MPL = \frac{w}{P}$ , here  $\frac{w}{P}$  is the real wage rate. At that point, the contribution of the last unit of labour will be just equal to output MPL. Thus one can conclude that the wages paid would be equal to the marginal product of labour.

Overall wages are determined by the market supply and demand for labour. The aggregate demand for labour reflects the  $MRP_L$ , but the supply curve is somewhat unique. There are a fixed number of hours in the week so a decision to supply labour is also a decision to demand leisure.

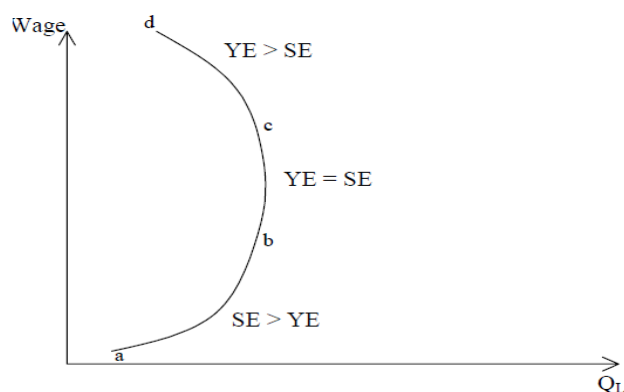
After deducting time for sleeping and eating, a person has 90 usable hours in a week. So a decision to work 40 hours is a decision to use 50 hours for other purposes. We can study the labour supply curve by examining the demand for leisure.

Consumers can buy back leisure time just as they can buy cars and groceries. Recall that any price change has two effects: an income effect and a substitution effect. These effects impact the demand for leisure and therefore the shape of the labour supply curve.

(1) **Substitution Effect (SE):** Consumers can buy their leisure by giving up an hourly wage. So the wage rate is the price (or opportunity cost) of leisure. When the wage rate rises, the price of leisure increases relative to other things a consumer can buy with a wage. So an increase in the wage leads to a reduction in the demand for leisure and an increase in the supply of labour.

(2) **Income Effect (YE):** Higher wages make consumers richer. The increase in wealth should lead to an increase in the demand for all goods, including leisure. Thus, a higher wage can lead to a reduction in the supply of labour.

Consider the labour supply curve drawn below.

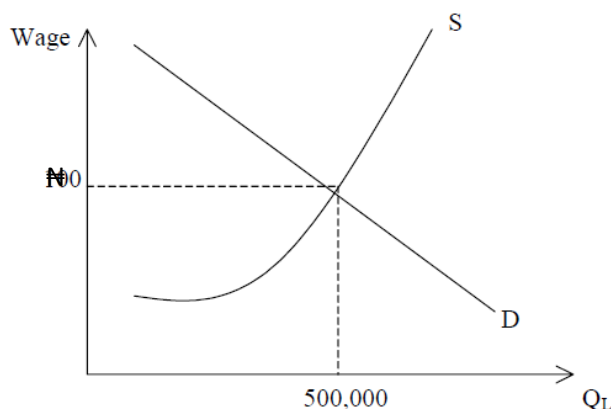


**Figure 3.15: Labour supply curve**

Depending on the worker, higher wages could lead to more, the same, or less labour supplied.

### 3.3 The Demand for Labour and Wage Determination

Profit-maximizing firms hire workers up to the point where the wage equals the  $MRP_L$ . Recall that the demand for labour is a derived demand. If product demand rises and increases the price of the good, then the  $MRP_L$  rises and the demand for labour shifts outward. Equilibrium occurs where the supply and demand for labour intersect. In the diagram below, the average daily wage is ₦100 and 500,000 workers are employed.



## Why Wages Differ

Of course, there is not one labour market. In fact, there are many of them. Each with its own supply and demand curves and equilibrium wages. In general, low-wages are found among the young, the disadvantaged, and the uneducated. Why are some wages so low, while others are so high?

Supply and demand tells us some of the story: wages are high when supply is low and demand is high, while wages are low when supply is high and demand is weak.

Why is the demand for labour greater in some markets than others?  $MRP_L$  is determined by the  $MPP_L$ , which depends on workers' own abilities and efforts on the job. Sometimes these characteristics are less important than the supply of other factors of production available for workers to work with. U.S. labour is more productive than Nigeria labour because U.S. workers have more machinery, natural resources, and technology to work with. As a consequence, manufacturing workers in the U.S. earn more than their Nigeria counterparts. Recall that the  $MRP_L$  rises with the amount of capital employed by the firm. The  $MPP_L$  is also influenced by education, training, and job experience.

The supply of labour differs across labour markets. The size of the available working population relative to the magnitude of industrial activity in an area is important. There are also non-monetary components (called compensating wage differentials) of any job that affects the wages paid and influences the supply of labour. Pleasant jobs such as teaching students in suburban schools attracts lots of labour so the wage is low. Unpleasant jobs, such as garbage collection, does not attract many people so the wages is fairly high. For example, professors at CSU campuses make 10% less than California prison guards. The prison guards receive a higher salary to compensate them for the dangers they face at work.

The amount of training needed to enter a profession will impact the supply of labour. Neurosurgeons and NBA stars have high incomes because there are few people as highly skilled as they and it is time consuming and expensive to acquire these skills, even for people with the ability to complete the necessary training.

### 3.3.1 Human Capital Theory

To go to college, you pay tuition, but you also forego earnings by attending school, and the loss of income is costlier. Your education is an investment in yourself. Like a firm that builds a plant to increase its future earnings, you are investing in the future hoping your college education will allow you to earn more in the future. You may also be able to find a job that is more prestigious or pleasant after graduating. Doctors and lawyers earn high salaries because of their many years of training. High wages are the return on their investment.

Jobs requiring more education pay higher wages to encourage people to go to school. Firms are willing to pay educated workers higher wages because they are more productive on the job.

Schooling and training raises the  $MRP_L$ . Universities are the factories that take unproductive workers as raw materials and apply training to produce workers that are productive as outputs.

This view of what happens in school makes educators happy and agrees with common sense.

However, some social scientists doubt that this is how schooling increases earnings power. The dissenting views are summarized below.

(1) **Education is a sorting mechanism:** Education doesn't teach students anything relevant to their jobs. People differ in ability before they enter school and leave the same way they came into school. Education sorts people by ability. More intelligent and disciplined people are successful in school and they will be more successful on the job. Harder working, more intelligent people stay in school longer and employers know this, so they hire educated workers.

(2) **Radical view:** Education doesn't sort people by ability; it sorts them by socio-economic class. The rich can buy better education and keep their kids in school regardless of ability. Education allows wealthy families to pass their economic status to their children, while making it appear valid for firms to give them higher salaries. Schools don't teach knowledge, they teach discipline: how to show up for work on time, how to speak respectfully, follow authority, etc. Businesses prefer these attributes and want to hire more educated workers. Schools teach docility and the acceptance of the capitalist status quo, making schooling attractive to firms.

(3) **Dual labour markets:** Good jobs are in the primary labour market. These jobs are interesting and offer career advancement. Education determines who will enter this market and there are financial rewards for those who go to school. Economists, who believe this theory, think that school does teach some valuable skills. But, they agree with the radicals --that wealthy kids have better access to schools and education.

Bad jobs (such as fast food or retail services) are in the secondary labour market. The jobs in this market are characterized by low-wages, few benefits, and no on-the-job training or advancement.

Lateness and absenteeism are frequent because the jobs are so bad that these workers develop bad work habits, which confirms the prejudice of those who assigned them to the bad jobs in the first place. Higher education doesn't lead to higher wages or greater benefits. So workers in the secondary labour market have little incentive to invest in education.

In summary, educated workers earn higher wages, but there are different beliefs about how education contributes to higher wages.

## **SELF-ASSESSMENT EXERCISE**

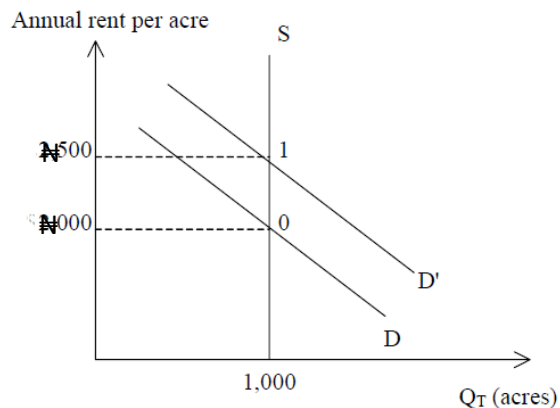
In your own words, how is wage determined in the labour market?

### **3.4 Determination of Rent**

This theory was first developed by the economist David Ricardo. It was called *the Ricardian Theory of Rent*. Ricardo defined rent as "that portion of produce of the earth which is paid to the landlord for the use of the original and indestructible powers of the soil". However, later on the modern theory of rent was developed by the modern economists. The main difference between the Ricardian Theory and this theory is that, Ricardian Theory used the difference between surpluses enjoyed from superior land to the inferior land. In the modern theory the rent is determined by the demand and supply

forces in the market just like the other factors of production. Demand for land means total land demanded by the economy as a whole. Demand for land like others depends upon the marginal revenue productivity. Rent paid by the economy will be equal to the marginal revenue productivity which is also subject to the law of diminishing returns. This suggests that the demand curve like any other demand curve will be downward sloping. It shows that the demand for land and rent are negatively related. On the other hand, supply of land for an economy is fixed, that is, it is perfectly inelastic.

Rent is the payment for use of land. The supply of land is unique --it is generally considered to be fixed. People can clear land, drain swamps, and fertilize it, but it's difficult to change the total supply of land with human effort. Consider the diagram below for a small town.



Notice the vertical supply --there are 1,000 acres of land in the town no matter what the rent is. The demand for land (T) reflects the  $MRP_T$ . Rent is determined by the intersection of supply and demand --₦2,000 per year. The level of rent is determined entirely by the demand side of the market. If the demand for land increases to  $D'$  as people move into the area, the annual rent rises to ₦2,500 per acre. The landowners earn more money, but they do not supply additional land.

An economic rent is a payment for a factor of production (e.g., land) that does not change the amount of the factor that is supplied.

### 3.4.1 Some Complications in the Rent of Land

All land is not identical --it differs with respect to soil, topography, access to sun and water, and proximity to the marketplace. Capital invested on any piece of land must yield the same return as capital invested on any other piece of land. For example, if a crop is produced for ₦160,000 (in labour, fertilizer, equipment, and fuel) a year on plot A and for ₦120,000 on plot B, then the rent on B must be ₦40,000 higher. Otherwise, the production on plot B is cheaper and this would lead farmers to bid for plot B until the rent was ₦40,000 higher than plot A.

Some pieces of land are such low quality that it doesn't pay to use them (e.g., remote deserts). Any land on the borderline between use and non-use is called marginal land. It earns no rent because if rent were charged, no one would use it.

Rent on any piece of land will equal the difference in the cost of producing output on that land and the cost of producing output on marginal land. If the population increases, the demand for land rises and (1) it will now pay to use some land whose former use was



unprofitable. For example, land that could not be given away in the west is now quite valuable. And (2) people will begin more intensive use of the land already in use.

For example, build high rise condos instead of single family homes. Land that is marginal after the change must be inferior to land that was marginal prior to the change. Thus, rents increase by the difference between the production costs on the old and new marginal lands. Consider the non-rent costs and rent on 3 pieces of land.

Type of Land (plot)	Non-land cost of producing crops	Total Rent	
		Before	After
A --better than marginal before and after	₦120,000	₦80,000	₦92,000
B --marginal before but not anymore	₦200,000	₦0	₦12,000
C --previously not worth using, now marginal	₦212,000	₦0	₦0

Crops cost ₦80,000 more when produced on B versus A, so the rent on A is ₦80,000. The increased demand for land is great enough to bring plot B into use. Plot C is now marginal land and B requires a rent of ₦12,000 --the cost advantage in using plot B over C in production. Plot A's rent must rise to ₦92,000 --its cost advantage over plot C (the new marginal land).

Rent also rises because of higher intensity in the use of land already in cultivation. As farmers apply more fertilizer and labour to the land, the  $MRP_T$  is rising and the landowners capture this in the form of higher rents. Remember the demand for land is driven by the  $MRP_T$ , so when the  $MRP_T$  rises, so do rents.

**Economic rent:** is any payment made to a factor above the amount necessary to induce any of that factor to be supplied to its present employment.

So a payment to an input has 2 components.

- (1) The minimum payment needed to acquire the input --to compensate labour for the loss of leisure, unpleasantness of work, etc. Otherwise, the worker won't supply any of his labour.
- (2) Bonus that goes to inputs of high quality --payments to talented workers or extra payments to a better piece of land. Understand that the bonus is great for the worker, but it isn't a deciding factor in the decision to work.

For land, the total supply of land is available for use whether rent is high, low, or zero. So payments to landowners are entirely economic rent, because they are unnecessary to induce the quantity of land in the economy.

### SELF-ASSESSMENT EXERCISE

“All land is not identical ...” how can rent be determined and what are the possible complications one can face?

### 3.5 Determination of Interest

The interest rate is the price at which funds are borrowed. Interest rates are determined by the supply and demand for loanable funds. Funds are loaned to users in many ways: mortgages, corporate or government bonds, and consumer credit. On the demand side are borrowers -- people or institutions that wish to spend more than they currently have. In business, loans are often used to finance investments.

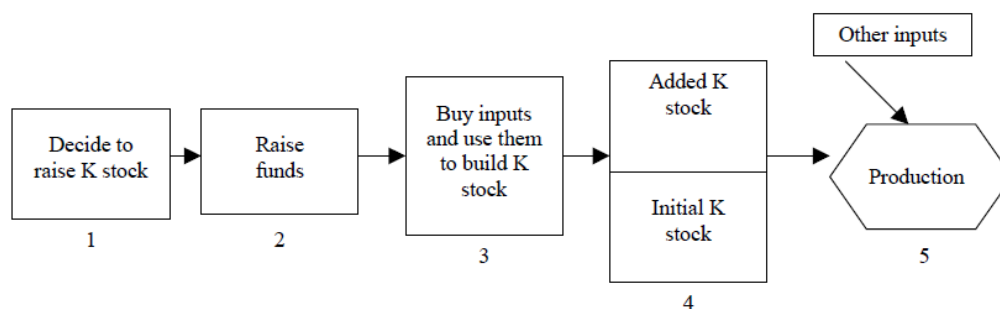
**Investment:** flow of resources into the production of new capital. It is the labour, steel, and other inputs devoted to the production of plants and equipment.

**Capital:** refers to a stock of plant, equipment, and other productive resources held by the firm, individual, or organization.

A firm borrows funds to finance an investment, which is the acquisition of the physical capital the firm will buy.

Investment is the rate at which capital grows. The larger the level of investment, the greater the rate at which capital grows. Think of filling a bathtub. The accumulated water in the tub is the stock of capital and the flow of water from the faucet is the flow of investment. The capital stock only increases when there is investment. If investment equals 0, then the capital stock remains constant and doesn't fall to zero.

#### 3.5.1 Investment Production Process



**Step 1:** firm decides to increase the capital stock.

**Step 2:** it raises funds to finance its expansion from outside sources, such as banks, or by holding onto part of its earnings, rather than paying them out as dividends to company owners.

**Step 3:** it uses the funds to hire inputs to build factories, warehouses, etc. This is the act of investment.

**Step 4:** the firm has a larger stock of capital once the investment is complete.

**Step 5:** capital is used (with other inputs) to increase production or lower costs. At this point, the firm begins earning returns on the investment.

If funds are borrowed, the firm must repay the lender. The payment is called interest and it is calculated as an annual percentage of the amount borrowed. For example, ₦1,000 is borrowed at 12% per year then the annual interest payment is ₦120.

Firms will demand the quantity of borrowed funds that makes the MRP of investment financed by the funds just equal to the interest payments.

Capital is different from any other input because it is durable and lasts for years. For example, a blast furnace (used in the production of steel) is a durable good because it contributes to today's production and to future production. Thus, calculating the  $MRP_K$  is more complex than other inputs.

To determine whether an investment is profitable (i.e.,  $MRP_K > \text{interest}$ ), requires a comparison of money values received at different times. To do this, economists use a procedure called "discounting".

It is important to know:

(1) A sum of money received in the future is worth less than the same sum of money received today. Compare ₦1 today versus ₦1 one year from today. If the interest rate is 10% then you could lend the ₦1 out for one year and earn ₦0.10 and receive ₦1.10 in one year. Thus, one naira received today is worth more than ₦1 received one year from today. In fact, the discounted value of ₦1 paid one year from now is only  $\text{₦}1/(1 + 0.10) = \text{₦}0.91$  today. And the discounted value of ₦1 received 2 years from now is  $\text{₦}1/(1 + 0.10)^2 = \text{₦}0.83$  today. In other words, if you placed ₦0.83 in the bank today and received a 10% annual interest rate, then in two years, you would have ₦1 in the bank. A little more work gives the formula for discounting a sum of money (₦X) paid  $n$  years from now =  $\text{₦}X/(1 + i)^n$  --where  $i$  is the interest rate.

(2) The difference in values of money today and money in the future is greater when the interest rate is higher. Now suppose the interest rate is 15%. Then ₦1.00 today yields ₦1.15 in one year (instead of ₦1.10). So the higher the interest rate, the greater the difference between a ₦1 received today and ₦1 in the future. When looking at the discounting formula  $[\text{₦}X/(1 + i)^n]$ , you can see that higher interest rates lower the value of ₦X paid in  $n$  years.

Interest rates are crucial in determining the economy's level of investment - or the amount of current consumption that people will forego to use resources to build machines and factories that can increase consumption in the future. The interest rate largely determines the allocation of society's resources between the present and future.

In the classical theory, supply and demand for capital determines the optimal interest rates. The rate of interest that is determined by the intersection of investment and saving is the price of investible resource (capital). The demand for capital is done by the entrepreneur for further investment and for the productive purpose.

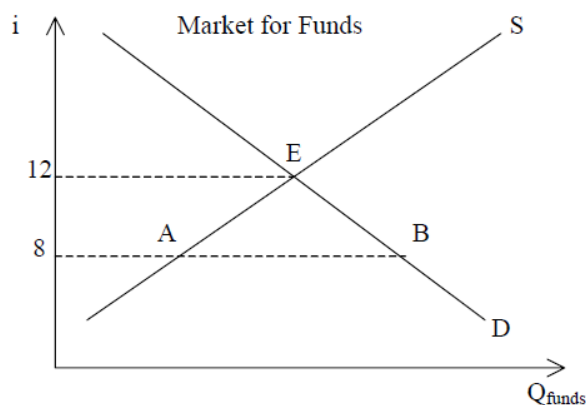
But the productivity of capital is dependent on the law of variable proportion. Which means as more and more of capital is employed the productivity from capital goes on decreasing. Therefore, the entrepreneur will employ only capital up to that level where the rate of interest is equal to the Marginal productivity of capital ( $MPK$ ).  $MPK = \frac{R}{P}$ , where  $MPK$  is the marginal productivity of capital and  $\frac{R}{P}$  is real rental for capital. It shows that demand for capital is inversely related to the rate of interest. There are many other factors which affect the demand for capital. On the other hand, supply of capital is positively related to the rate of interest. As the rate of interest increases the savings increases and vice versa. Thus the entrepreneur will employ the capital where  $MPK$  is equal to real price of capital. The optimal rate of interest rate is determined by the intersection of demand and supply curves. If the rate of interest rises above the equilibrium interest rates the demand for the investment will decline and the supply of

savings will increase. As there is excess of savings than demand in the economy the market forces will bring the interest rate to the original interest rates. Now, if the interest rates fall below the optimal level then there is excess of demand than supply in the economy and hence, the market forces will adjust the interest rates.

### 3.5.2 The Demand and Supply of Funds

The quantity of loanable funds demanded will fall as interest rates rise. The demand for borrowed funds is derived from the desire to invest in capital goods. But, the  $MRP_K$  is received in the future. Thus, the  $MRP_K$  in terms of today's naira shrinks as the interest rate rises. Why? Because future returns must be discounted when the interest rate rises. So a machine that looks like a good investment at a 10% rate of interest may look terrible at 15%. So higher interest rates lower the demand for capital.

Loans look better to lenders when they receive higher interest rates, so the supply of loanable funds is upward sloping. The equilibrium interest rate occurs where the supply and demand curves intersect. This is shown below.



**Figure 3.16: Equilibrium Interest rate**

Are interest rates too high? Usury laws try to restrict the interest rate. Black markets often develop and people who are willing to lend charge rates of interest that are higher than the equilibrium rate to compensate them for the risk of being caught. If a usury law prohibits interest rates above 8% then the quantity demanded is greater than the quantity supplied. So applicants for bank loans are turned down even if they have good credit. The people who gain are the lucky consumers who get the loans at the cheaper interest rate. Banks who want to lend at 12% and borrowers who want funds at 12% lose. Few people sympathize with bank stockholders and the consumers who get the loans at 8% are happy. The unlucky consumers don't blame the government's usury law. Usury laws have no effect if the limit on interest is 8% and the equilibrium is 6%, but the law does have an effect if the equilibrium interest rate is 12%.

### SELF-ASSESSMENT EXERCISE

Explain how interest rate is determined.

### 3.6 Determination of profit

Profit is another important factor in factor payments. This theory was first developed by Edgeworth, Chapman, Stigler, and Stonier. This theory is also dependent upon the marginal revenue productivity. It is also called marginal product and capital demand.

Let one consider an example. The main objective of firm is to maximize profit. As we know that profit would be difference between the revenue and costs  $Profit = PY - WL - RK$ . Where the revenue would be equal to the price of the good multiplied by the output of the firm  $P \cdot Y$ . On the other hand, the costs of the firm include labour costs  $WL$ , capital costs  $RK$ , rent cost if any. Now if we substitute our production function  $Y = f(L, K)$ . Then we would see that the profit of the firm is depended on factor prices and factor inputs  $Profit = Pf(L, K) - WL - RK$ . Hence firm would choose the optimal level of factor inputs that would maximize profit of the firm.

Are profits too high or too low?

Economists are unlikely to state that factor prices are "too high" or "too low" in some moral or ethical sense. They are more likely to ask: what is the market equilibrium price and then ask whether there are any good reasons to interfere with the market equilibrium. This analysis isn't easy to apply to profits when you are uncertain as to which factor of production earns profits.

Profits are a residual --they are what remain from the selling price after all the other factors have been paid. So what factor of production receives this reward? What factor's MRP constitutes the profit rate?

Economic profit is the amount the firm earns over and above the payments for all the inputs, including the interest paid on any capital supplied by the firm's owners. Profit is the payment of the opportunity cost of capital the firm's owners provide to the firm. Profit is closely related to the interest rate. In a pretend world where everything is certain and unchanging, the capitalists who invested money in firms would simply earn the market rate of return on their funds. Profits above the market interest rate would be competed away and profits below this level would causes capitalists to withdraw their funds from firms and deposit them into banks. In this pretend world, capitalists are just moneylenders.

But, in the real world, capitalists are more than moneylenders and they often earn returns that exceed the interest rate. Active capitalists that seek out and create earnings opportunities are called entrepreneurs. There are 3 primary ways in which entrepreneurs are able to drive profits above interest levels.

(1) **Exercise of Monopoly Power:** If an entrepreneur can establish a monopoly over some or all of her products, even for a short time, she can use the monopoly power of her firm to earn profits.

(2) **Risk Bearing:** An entrepreneur can engage in risky activities. For example, when a firm prospect for oil, it drills an exploratory well hoping to find a pool of petroleum at the bottom. Yet, a high proportion of such attempts produce only dry holes and the cost of the work is wasted. Of course, if the investor is lucky and finds oil, she is rewarded handsomely. The income she receives is a payment for taking risk. A few lucky people make out well, but most suffer large losses. How well do we expect risk takers to do on average? If 1 out of every 10 exploratory drilling pays off, we expect its return to be 10 times as high as the interest rate.

(3) **Returns to Innovation:** This is most important from a social welfare standpoint. The entrepreneur who is first to make a desirable new product or use a cost-saving machine will receive a higher profit than that going to a less innovative (but similar)

business manager. Innovation is not the same as invention. Invention is the act of generating a new idea and innovation is the act of putting the new idea into practical use. A person who is able to bring a new product to market or produce a product more cheaply will temporarily be ahead of the competition by selling to more customers.

Monopoly profits --the reward for innovation -- are only temporary because once other firms see that the idea is profitable, they will imitate it. Even if they can't produce the same exact product, they must find close substitutes in order to survive. The innovator's monopoly position comes to an end over time with competition. So to keep making profits above market returns, she must come up with new innovations.

The growth of the capitalist system is fuelled by entrepreneurs' search for new ideas in order to earn profits.

### **3.6.1 Profit Taxation**

Profit above the market rate of return is considered to be the return to entrepreneurial talent. But, no one knows exactly what entrepreneurial talent is, because you can't measure it or teach it. So we don't know how the observed profit rate relates to the minimum reward necessary to attract entrepreneurial talent into the market. Danger: if we tax away profits, then this may cause fewer people to enter the market in search of innovations and profit.

For example, if oil companies' profits are mostly comprised of rents, then we could tax away the rents without reducing oil production and exploration. But, profits may not be largely economic rent. So increasing profit taxes may drain the lifeblood out of the capitalist system.

## **4.0 CONCLUSION**

In this unit we have seen that wages—the price of labour—are determined (like all prices) by supply and demand. When workers sell their labour, the price they can charge is influenced by several factors on the supply side and several factors on the demand side. The most basic of these is the number of workers available (supply) and the number of workers needed (demand). In addition, wage levels are shaped by the skill sets workers bring and employers need, as well as the location of the jobs being offered.

However, it is important to avoid confusing rent which refers to factor payment in excess of transfer earning and rent which refers to the payment made by a tenant to a landlord for the hiring of land and/or a building. Rent may be earned by any factor. High rent is a result, not a cause, because demand for land is a derived demand. The higher the price of the product produced by the use of the land, the greater will be the demand for land. Hence, price is not high because a rent is paid, but a rent is paid because price is high.

Furthermore, profit is the net amount after adding up all the revenue and deducting all the expenses, costs and taxes. If expenses surpass the revenue, then the business makes a loss.

## **5.0 SUMMARY**

In this unit, we saw how the interplay between all of these factors will eventually cause wages to settle—that is, the number of workers, the number of jobs, the skills involved,

and the location of the jobs will eventually lead workers and employers to reach a series of wage agreements. If employers (demand) cannot find enough workers to meet their needs, they will keep raising their wage offers until more workers are attracted. If workers are in abundance (supply), wages will fall until the surplus labour decides to go elsewhere in search of jobs. When supply and demand meet, the equilibrium wage rate is established. Also rent has the function of allocating the factor to the highest-valued competing uses. It reveals which uses are the highest valued and directs the factor to those uses.

## **6.0 TUTOR-MARKED ASSIGNMENT**

Some occupations that do not have pleasant working conditions, such as rubbish collection, receive low pay, while those with pleasant conditions, such as senior managers, receive high pay. How far does economic analysis explain this situation?

## **7.0 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. (2001). *Modern Microeconomics*, (2nd ed.). Hampshire: Macmillan Press Ltd.

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

## **MODULE 3: THEORIES OF FIRM, DISTRIBUTION AND COST BENEFIT ANALYSIS**

- Unit 1: Theory of The Firm**  
**Unit 2: Theory of Distribution**  
**Unit 3: Cost-Benefit Analysis**

### **UNIT 1: THEORY OF THE FIRM**

#### **CONTENTS**

- 1.0 Introduction
- 2.0 Objectives
- 3.0 Main Content
  - 3.1 Concept of Theory of the Firm
  - 3.2 Profit-Maximizing Theories:
    - 3.2.1 Simple Mathematics of the Profit Maximization Hypothesis:
    - 3.2.2 Criticisms of Marginalist Theory of the Firm:
    - 3.2.3 Criticisms of the Modern Approach:
  - 3.3 Other Optimizing Theories:
    - 3.3.1 Rationale:
    - 3.3.2 Implications and Limitations:
    - 3.3.3 Mathematical Presentation:
    - 3.3.4 The Dynamic Model:
  - 3.4 Non-Optimizing Theories:
    - 3.4.1 The Behavioural Theory of the Firm:
    - 3.4.2 Goals of the Firm:
    - 3.4.3 Cyert and March Organizational Slack
- 4.0 Conclusion
- 5.0 Summary
- 6.0 Tutor-Marked Assignment
- 7.0 References/Further Reading

#### **1.0 INTRODUCTION**

The theory of the firm is the microeconomic concept founded in neoclassical economics that states that a firm exists and make decisions to maximize profits. The theory holds that the overall nature of companies is to maximize profits meaning to create as much of a gap between revenue and costs. The firm's goal is to determine pricing and demand within the market and allocate resources to maximize net profits.

#### **2.0 OBJECTIVES**

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

- ✓ Understand the traditional and modern Profit-Maximizing Theories
- ✓ Explain Baumol's Single Period Sales (Revenue) Maximization subject to Profit Constraint
- ✓ Explain the Behavioural Theory of the Firm
- ✓ Understand the goals of the Firm



### **3.0 MAIN CONTENT**

#### **3.1 Concept of Theory of the Firm**

In the theory of the firm, the behaviour of any company is said to be driven by profit maximization. The theory governs decision making in a variety of areas including resource allocation, production techniques, pricing adjustments, and the volume of production.

Early economic analysis focused on broad industries, but as the 19th century progressed, more economists began to ask basic questions about why companies produce what they produce and what motivates their choices when allocating capital and labour.

Under the theory of the firm, the company's sole purpose or goal is to maximize profit. However, the theory has been debated and expanded to consider whether a company's goal is to maximize profits in the short-term or long-term.

Modern takes on the theory of the firm sometimes distinguish between long-run motivations, such as sustainability, and short-run motivations, such as profit maximization. The theory has been debated by supporters and critics.

If a company's goal is to maximize short-term profits, it might find ways to boost revenue and reduce costs. However, companies that utilize fixed assets like equipment would ultimately need to make capital investments to ensure the company is profitable in the long-term. The use of cash to invest in assets would undoubtedly hurt short-term profits but would help with the long-term viability of the company.

Competition can also impact the decision making of company executives. If competition is strong, the company will need to not only maximize profits but also stay one step ahead of its competitors by reinventing itself and adapting its offerings. Therefore, long-term profits could only be maximized if there's a balance between short-term profits and investing in the future.

The theory of the firm supports the notion that profit maximization is the nature of a company's existence, but today companies must consider shareholder wealth through dividends, public perception, social responsibility, and long-term investments in the company's viability.

Risks exist for companies that subscribe to the profit maximization goal as stated under the theory of the firm. Solely focusing on profit maximization comes with a level of risk in regards to public perception and a loss of goodwill between the company, consumers, investors, and the public.

A modern take on the theory of the firm proposes that maximizing profits is not the only driving goal of a company particularly with publicly held companies. Companies that have issued equity or sold stock have diluted their ownership. The low equity ownership by the decision makers in the company can lead to chief executive officers (CEOs) having multiple goals including profit maximization, sales maximization, public relations, and market share

Further risks exist when a firm focuses on a single strategy within the marketplace to maximize profits. If a company relies on the sale of one particular good for its overall success, and the associated product eventually fails within the marketplace, the

company can fall into financial hardship. Competition and the lack of investment in its long-term success such as updating and expanding product offerings can eventually drive a company into bankruptcy.

### 3.2 Profit-Maximizing Theories

The traditional objective of the business firm is profit-maximization. The theories based on the objective of profit maximization are derived from the neo-classical marginalist theory of the firm.

In essence the theories based on the profit- maximization goal suggests that firm seeks to make the difference between total revenue (or sales receipt) and total cost (outgo) as large as possible.

However, one pertinent question here is: does the firm attempt to maximize long term profit or short- term profit? The basic valuation model of the firm is based on the fundamental assumption that the firm seeks to maximize its long-term profit.

According to this model, a firm seeks to maximize its discounted present value. To arrive at an estimate of discounted present value of the firm we reduce future profits by a discount factor or weight, to make future profits comparable with present profits. Let  $PV_f$  refer to the present value of the firm and  $\pi_1, \pi_2, \dots, \pi_n$  refer to profits in the next  $n$  time periods. Therefore, we can express  $PV_f$  as:

$$PV_f = W_1 \pi_1 + W_2 \pi_2 + \dots + W_n \pi_n$$

where  $W_1, W_2, \dots, W_n$  are the weights we assign to future profits to be able to make inter-temporal comparisons of money sums. One complication that arises in this context is that the choice of weights largely depends on the firm's rate of time preference, i.e., how the firm values present profits compared to future profits.

The short-run profit maximization hypothesis is based on the famous marginalist rule which we have explained. A firm maximizes profit when by producing and selling one more unit it adds as much as to revenue as to cost.

The addition to revenue is called marginal revenue and the additional cost marginal cost. Thus, a firm maximizes profit when  $MR = MC$ . If this condition holds and if the MC curve intersects the MR curve from below and not from above, total profit (i.e.,  $\pi = TR - TC$ ) will be maximum.

However, if the periods are dependent (i.e., if current decisions or actions affect future decisions of the firm) short-run profit maximization will lead to incorrect decisions because of lack of provision for the future. For instance, the firm could generate higher profits now by not replacing capital goods, delaying payment on due accounts etc. all of which will surely reduce the size of future profits.

By contrast, if profits are independent in different time periods, long-run profit maximization would simply amount to maximizing the series of short-term profits. But such a situation does not prevail in the real world. All firms which have made huge capital investments will observe that profits in different time periods are interdependent.

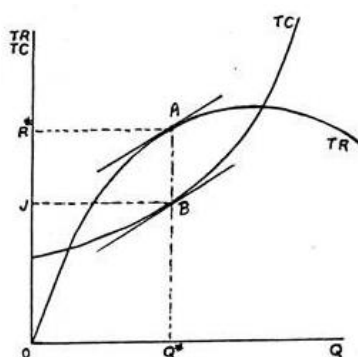
There is a trade-off between short-term and long-term profit. If more (or a steady flow of) profit is derived in the long run, adequate provision has to be made for depreciation

(capital consumption) and short-term dues are to be cleared. If more profit is to be made in the short run, some long-term profit has to be sacrificed.

With the above complications in mind we may now briefly discuss the traditional theory. The essence of the traditional approach is to compare cost and revenue of a firm at different levels of output and to select the one which maximizes the absolute differences between the two.

The short-run profit maximization hypothesis is illustrated in Figure 1. The TC and TR are shown on the vertical axis and output on the horizontal axis. The firm produces a level of output  $OQ^*$  for which  $TR = OR^*$  and  $TC = OJ$  and the gap between the two ( $R^*J$ ) is maximum. Thus  $Q^*$  is indeed the profit-maximizing level of output.

The slope of the TR curve measures MR and the slope of the TC curve measures MC. At points A and B, two curves have the same slope. Thus at  $OQ^*$ ,  $MR = MC$ . This can be verified by passing two tangents — one through A and the other through B and ensuring that they are parallel.



### ***Single Period Profit Maximization***

The total cost curve is always non-linear and has got nothing to do with the market structure. The slope of the revenue curve depends on elasticity of demand and is crucially dependent on the market structure. Since most real life markets are imperfectly competitive we assume non-linear total revenue function, too.

Subtracting the TC curve from the TR curve we derive the total net profit curve  $\pi$  which cuts the horizontal axis where  $TR = TC$ . We reach the top of the profit hill when  $Q^*$  is the level of output that is produced and sold.

In Figure 1 the firm produces  $OQ^*$  units and makes a total revenue of  $OR^*$  by charging a price of  $OR^*/OQ^*$ . At this stage total profit is  $R^*J$  which is maximum.

The hypothesis is based on a number of assumptions. Prima facie, the decision-maker (manager or entrepreneur) is supposed to have relevant information about cost and revenue on the basis of which an optimal decision can be made. Secondly, he is assumed to have sufficient power to make a decision and implement it properly.

However, the external or market forces — which are beyond the control of a firm or its management — are the major determinants of the firm's optimal decision on price and quantity. This theory is universally applicable.

#### **3.2.1 Simple Mathematics of the Profit Maximization Hypothesis:**

The equilibrium of the profit-maximizing firm occurs simultaneously on the input and output sides — i.e., a firm which is maximising its profit by choosing an output at which

marginal cost equals marginal revenue is simultaneously minimizing the cost of producing that output, or maximizing the output subject to cost constraint.

We can now prove that minimizing the cost of the prescribed level of output requires satisfaction of the same condition as does maximization of the output subject to cost constraint. So the latter condition is also a condition for profit maximization.

**Minimizing the Cost of the Prescribed Level of Output:**

$$\text{Let } Q = f(K, L) \dots \dots \dots (1)$$

be the production function, where Q is output and K and L are the quantities of two types of factor services.

$$\text{Let } Q_0 = f(K, L) \dots \dots \dots (2)$$

be the prescribed output, and

$$C = rK + wL \dots \dots \dots (3)$$

where C is total cost and r and w are the factor-service prices of K and L, respectively.

Then, in order to minimize (3) subject to (2), form the function

$$S = rK + wL + \lambda[Q_0 - f(K, L)] \dots \dots \dots (4)$$

$$\frac{\partial S}{\partial K} = r - \lambda \frac{\partial Q}{\partial K} = 0$$

$$\frac{\partial S}{\partial L} = w - \lambda \frac{\partial Q}{\partial L} = 0$$

$$\frac{\partial S}{\partial \lambda} = Q_0 - f(K, L) = 0$$

Yielding as a condition of equilibrium

$$\frac{\frac{\partial Q}{\partial K}}{\frac{\partial Q}{\partial L}} = \frac{r}{w}$$

The second-order conditions (not shown) require that the isoquants be convex to the origin.

**Maximizing the Output Subject to Cost Constraints:**

Given equations (1) and (3) above, let the prescribed cost outlay be

$$C_0 = rK + wL \dots \dots \dots (5)$$

Then, in order to maximize (1) subject to (5) from the equation

$$M = f(K, L) + \lambda(C_0 - rk - wL) \dots \dots \dots (6)$$

$$\frac{\partial M}{\partial K} = \frac{\partial Q}{\partial K} - r = 0$$

$$\frac{\partial M}{\partial L} = \frac{\partial Q}{\partial L} - w = 0$$

$$\frac{\partial M}{\partial \lambda} = C_0 - rk - wL = 0$$

Yielding as a condition of equilibrium

$$\frac{\frac{\partial Q}{\partial K}}{\frac{\partial Q}{\partial L}} = \frac{r}{w}$$

The second-order conditions remain the same

### Maximizing Profit:

In the case of pure competition, let the price of the goods be represented by  $p$  and profit by  $\pi$ . Then, from (1) and (3),

$$\pi = pQ - C \dots \dots \dots (7)$$

Substituting from equation (1) and (3) for  $Q$  and  $C$  respectively,

$$\pi = p \cdot f(K, L) - rK + wL \dots \dots \dots (8)$$

Maximising  $\pi$

$$\frac{\partial \pi}{\partial K} = p \frac{\partial Q}{\partial K} - r = 0 \text{ or, } p \frac{\partial Q}{\partial K} = r$$

The value of the marginal product equals the factor-service price

$$\frac{\partial \pi}{\partial L} = p \frac{\partial Q}{\partial L} - w = 0 \text{ or, } p \frac{\partial Q}{\partial L} = w$$

$$\frac{\frac{\partial Q}{\partial K}}{\frac{\partial Q}{\partial L}} = \frac{r}{w} \text{ (a condition for profit maximisation assuming the second order condition is to be met.)}$$

### Why Maximum Profit?

From the above hypothesis we may provide two important rationale for maximizing profit.

Firstly, in a single owner firm, where the entrepreneur is both owner and manager, maximizing profit will maximize his own income. For a given amount of effort this is considered to be rational behaviour, irrespective of the structure of the market (or nature of competition).

If, however, the magnitude of profit varies with the amount of entrepreneurial effort expended, and effort has negative utility (disutility) for the entrepreneur, rational behaviour would dictate something else. He must find an optimal trade-off between effort and profit to maximize entrepreneurial utility which is unlikely to lead to maximum profit.

Secondly the impact of competition from rival firms forces the entrepreneur to maximize profits. Profit maximization therefore is not an aspect of discretionary behaviour (choice) but rather a compelled necessity. The entrepreneur is forced to maximize profit for his long-term survival.

Thus, the justification for profit maximization depends upon the nature of competition. If competition is absent (as in monopoly) there is no such pressure, although the previous argument still holds. Under highly competitive conditions the entrepreneur has to maximize profit just for survival.

### **3.2.2 Criticisms of Marginalist Theory of the Firm**

The profit maximization hypothesis developed during 1874-1890 by Leon Walras, W. S. Jevons and Alfred Marshall has formed the basis of the neo-classical (marginalist) theory of the firm. It has not been challenged up to the 1920's. But from early 1930s it has been subject to various criticisms.

Critics have argued that profit maximization is not the only objective of a firm. Modern business firms and their managers pursue certain other goals, too. Thus profit-maximization as the only goal of a firm is no longer a tenable hypothesis.

Being dissatisfied with both of the justifications, modern economists and management specialists have suggested various alternatives to profit-maximization.

The following arguments bear relevance in this context:

#### **1. Emergence of oligopoly:**

In the interwar period it became increasingly apparent, especially in industrially advanced countries, that a modern economy was dominated by oligopoly, a market structure characterized by the existence of a few large firms.

In a number of industries the structure has become gradually more and more concentrated (through merger or amalgamation) so that a few large (and dominant) firms accounted for a major portion of an industry's output.

In such environments there was hardly any pressure on each firm to maximize profit independently. Instead firms arrived at joint profit maximization through such devices as collusions and cartels. Alternatively put, the pressure from rival producers was not strong enough to dictate profit maximization as an inevitable objective for each firm.

#### **2. Separation of ownership from management:**

Secondly, in 1932, Berle and Means challenged, through their pioneering work, the argument that the firm would seek to maximize profits (even though it was necessary due to competitive pressure).

They discarded profit-maximization as a rational behaviour because of an alleged break in the identity of purpose of the manager and his firm. They discovered that in most large U.S. companies there was separation of ownership from control.

Most of such corporations were essentially in the control of the managers rather than the owners (shareholders), due to fragmentation and dispersion of ownership of shares.

Thus, in a handful of cases could a small group of shareholders' directly affect the decisions of the corporation. In such a situation, with managers acquiring only new

shares, the identity of purpose of maximizing profits and maximizing entrepreneurial satisfaction was largely shattered.

In truth, the notion of the entrepreneur has lost relevancy with management becoming an executive function performed by a committee, rather than a simple individual taking all decisions unilaterally (as has been postulated — explicitly and implicitly — by the traditional marginalist approach).

The inevitable consequence of such divorce of ownership from control was that managers may wish to pursue goals other than profit maximization, and would be forced to take into consideration the matter of profits to the extent that sufficient cash had to be generated to pay satisfactory dividends to the shareholders (so that they did not withdraw funds from the company).

### **3.2.3 Criticisms of the Modern Approach:**

Although this view has been accepted by many modern economists, the trend towards this type of change in power is not universal. Supporters of the traditional viewpoint would argue that the share-holders have ultimate power and, if properly motivated, can exert considerable influence.

At times, at the annual general meeting of a company, shareholders are able to put a lot of pressure on managerial decisions. Secondly, it has been argued that an increase in the number of firms does not necessarily imply growing competition.

There may be keen competition among 3 to 4 dominant firms in an industry. Thus the need for making maximum profit is not stronger under pure competition than under oligopoly.

Those who believe that the profit-maximization is no longer a tenable hypothesis have suggested a number of alternatives.

These fall into two broad categories:

- (1) Those who hold that something else other than profit is maximized and
- (2) Those who postulate non-maximizing behaviour.

### **SELF-ASSESSMENT EXERCISE**

Briefly explain the neo-classical marginalist theory of the firm and its criticism

### **3.3 Other Optimizing Theories:**

There are various alternative approaches to profit maximization. Here we restrict ourselves to the most important ones.

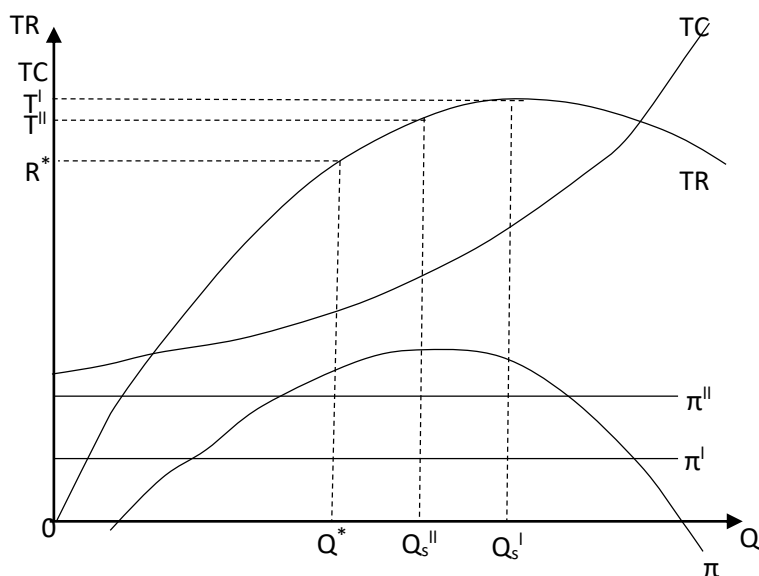
Baumol's Single Period Sales (Revenue) Maximization subject to Profit Constraint:

One alternative to profit maximization has been suggested by W.J. Baumol that firms operating in oligopoly will seek to maximize sales revenue subject to a profit constraint.

His argument is largely, if not entirely, based on "public statements by businessmen and on a number of a priori arguments as to the disadvantages of declining sales, for example, fear of customers shunning a less popular product, less favourable treatment

from banks, loss of distributors and a poorer ability to adopt a counter strategy against a competitor.”

Baumol’s basic argument is summarized in Figure 3.17, which enables us to understand the difference between profit maximization and sales maximization.



**Figure 3.17: Baumol’s Sale-Maximizing model**

Total profit is maximized when the firm produces  $OQ^*$  units of output (as in Figure 1) Sales maximization, on the other hand, refers to maximization of total revenue ( $\pi = P \times Q$ ), rather than maximization of  $\pi$  (It is because if a firm quotes zero price it can sell an astronomical amount but its total revenue will be zero.) Total revenue is maximum when  $MR = 0$ , and  $MR = 0$  when the demand for a company’s product is unitary elastic.

In Figure 2 we observed that if the firm wishes to maximize total revenue (without profit constraint) it will choose output  $Q^s$ , where TR is maximum (i.e., the slope of the TR curve is zero or  $MR = 0$ ). However, Baumol has argued that, a constraint operates from shareholders. They require a minimum sum as dividend which would keep them content.

Alternatively put, shareholder demand a level of absolute profit of some amount which is exogenous (i.e., determined outside the model). If this minimum acceptable level of profit were  $\pi'$ , the firm could produce  $Q^s$  and still generate profits greater than  $\pi'$ . Hence in this situation it will be worthwhile to produce  $Q^s$ .

Likewise if the minimum acceptable profit is  $\pi''$ ,  $Q^s$  will not generate sufficient profits. The firm will have to reduce output to  $Q^*$  which is indeed the optimal output with the profit constraint specified.

Baumol’s model thus predicts that profits will be sacrificed for revenue. The sales maximizing level of output will exceed the profit-maximizing level and can only be sold at a lower price under imperfectly competitive market conditions.



In fact, the first main difference between the profit maximizer and a constrained sales maximizer is that the latter can charge a lower price to sell the extra ( $OQ''s - OQ^*$ ) output. This has to be the case if both have the same demand (AR) curve.

In terms of Figure 2, the profit maximizer produces  $OQ^*$  and charges a price of  $OR^*/OQ^*$  (= total revenue + output). Alternatively, the sales maximizer produces (in the  $\pi''$  constrained case)  $Q''s$  and sells at a price of  $OT''/OQ''s$ .

### **3.3.1 Rationale:**

Baumol's model no doubt carries enormous good sense. The motivation to maximize sales revenue is justified on the ground that the managers of large firms stand to gain more from this strategy than from profit maximization. Sales maximization implies expanding the size of the organization, enhancing the status of managers as also their promotion prospects.

Again their wages and compensation are directly related to responsibility, which, in its turn, is again an increasing function of size. Conversely, as Baumol argues, it is quite irrational for managers to maximize profits for shareholders when they will get hardly anything themselves. (It is just 'head I win, tail you lose' type of affair — one-sided game, that is).

### **3.3.2 Implications and Limitations:**

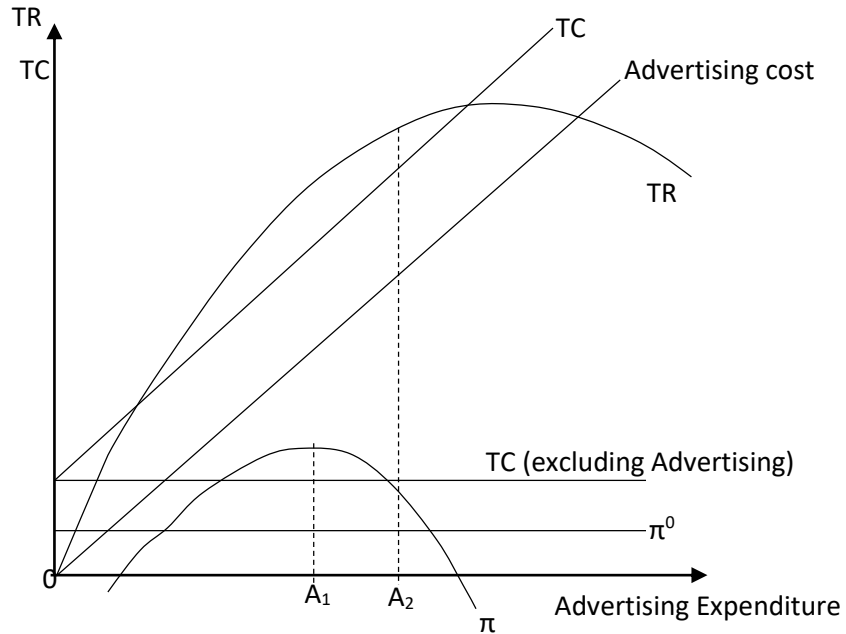
Baumol's' model is a single-period sales maximizing model. It applies at a single moment of time — i.e., it is static in nature. However, the model can be made dynamic for an in-depth study of multi-period optimization.

For this it will be necessary to consider various combinations of sales and revenues over time. In that case profit would be endogenous (i.e., determined from within the model) and would form the vehicle for growth through reinvestment of funds. This would enable us to predict an optimal combination of profits and growth rate of revenue. Such a dynamic model is appended below.

### **With Advertising:**

Secondly, advertising has been integrated into Baumol's model with consequent effect on the total revenue curve. Baumol's model has the implication that the sales-maximizing firm will spend more on advertising than the profit-maximizing firm.

Here Baumol simply assumes that advertising does not affect the market price of the product. But it leads to increase in the volume of sales (with diminishing returns). Hence it is assumed that advertising will always lead to a rise in TR, i.e., MR will never be negative. Baumol's extended model is illustrated in Figure 3.18.



**Fig. 3.18: Sales maximization with Advertisement**

Here the TC line is derived on the basis of the assumption that advertising (selling cost) does not affect total non-advertising cost. Now we measure advertising expenditure on the horizontal axis, and profit, revenue and cost on the vertical axis.

The TC curve is derived by superimposing the curve showing advertising cost, on the original TC (excluding advertising curve). Since there is positive (though not perfect) correlation between TR and advertising expenditure, the TR curve is upward sloping (its slope is positive though diminishing after a certain point due to diminishing returns).

Since advertising will always increase TR, the businessman will go on increasing advertising expenditure until prevented by the profit constraint.

In Baumol's model, therefore,  $A_1$  will be the profit maximizing level of advertising expenditure, which, if falls short of maximum profits, will invariably be less than the constrained maximizer's expenditure  $A_2$ .

Baumol's model, however, is not free from defects. It is inconsistent in one point at least. If advertising leads to greater output sold, non-advertising costs would be expected to rise. Yet, Baumol, in his simplified model, assumed that they would not.

### 3.3.3 Mathematical Presentation:

We may now seek to present a mathematical model on the line of Baumol.

We start by defining some important relations:

$$\frac{d(TR)}{dA} > 0, \frac{d(TC)}{dQ} > 0 \text{ and } Q > 0$$

The Lagrange function may now be formed as

$$L = R(Q, A) + \lambda(TR - TC - A - \bar{\pi})$$

The necessary conditions for a maximum are:

$$\frac{\partial L}{\partial Q} \leq 0, \quad \frac{\partial L}{\partial A} \leq 0, \quad \frac{\partial L}{\partial \lambda} \geq 0$$

Now differentiating the L function with respect to Q we get

$$\frac{\partial L}{\partial Q} = \frac{\partial R}{\partial Q} + \lambda \left[ \frac{R}{Q} - \frac{C}{Q} \right] \leq 0$$

So long  $Q > 0$ , the above expression holds as an equality. Solve for  $\frac{\partial C}{\partial Q}$

$$\frac{\partial C}{\partial Q} = \frac{\lambda + 1}{\lambda} * \frac{\partial R}{\partial Q} = \left( 1 + \frac{1}{\lambda} \right) \frac{\partial R}{\partial Q}$$

Since  $\lambda > 0$  it is obvious that  $\frac{\partial C}{\partial Q} > \frac{\partial R}{\partial Q}$  which implies that  $MC > MR$  in this model.

On the contrary, in case of a traditional profit maximising firm,  $MC = MR$ . likewise differentiating with respect to A, we get

$$\frac{\partial L}{\partial A} = \frac{\partial R}{\partial A} + \lambda \left[ \frac{\partial R}{\partial A} - \frac{\partial(AE)}{\partial A} \right] \leq 0$$

Since  $A > 0$  the above equation holds as an equality. Now let us solve for  $\frac{\partial(AE)}{\partial A}$

$$\frac{\partial(AE)}{\partial A} = \left( 1 + \frac{1}{\lambda} \right) \frac{\partial R}{\partial A} \text{ or } \frac{\partial(AE)}{\partial A} / \frac{\partial R}{\partial A} = \left( 1 + \frac{1}{\lambda} \right)$$

In Baumol's scheme for a profit maximiser, this ratio is always equal to 1.

As long as  $\lambda > 0$ , the advertising expenditure will be higher for a sales maximising firm. Thus, for a sales-revenue maximising firm, we arrive at the following

$$\left[ \frac{\partial C}{\partial Q} / \frac{\partial R}{\partial Q} \right] = \left[ \frac{\partial(AE)}{\partial A} / \frac{\partial R}{\partial A} \right] = \left( 1 + \frac{1}{\lambda} \right)$$

The implication is that excess profit or surplus will be partly utilized for advertising and partly for enhancing production. Baumol's model can be generalised with respect to multi-product firm. The product-mix of a revenue-maximizer will not be the same as that of a profit-maximizer.

**Illustration:** Given the demand function  $P = 20 - Q$  and the total cost function  $C = Q^2 + 8Q + 2$ , answer the following questions:

- What output,  $Q_\pi$ , maximizes total profit and what are the corresponding values of price,  $P_\pi$ , profit,  $\pi_\pi$ , and total revenue (sales),  $R_\pi$ ?
- What output,  $Q_r$ , maximizes sales and what are the corresponding values of price  $P_r$ , profit  $\pi_r$ , and total revenue,  $R_r$ ?

**Solution:**

(a) Now total profit =  $\pi = PQ - C$   
 $= -Q^2 + 20Q - Q^2 - 8Q - 2 = -2Q^2 + 12Q - 2$   
 In order to maximize profit, we require

$$\frac{d\pi}{dQ} = -4Q + 12 = 0 \quad \text{or} \quad Q_{\pi} = 3.$$

and so

$$P_{\pi} = 20 - Q_{\pi} = 17$$

$$\Pi_{\pi} = -2Q_{\pi}^2 + 12Q_{\pi} - 2 = -18 + 36 - 2 = 16$$

$$R_{\pi} = Q_{\pi}P_{\pi} = 3 \times 17 = 51$$

(b) Here total revenue =  $R = PQ = -Q^2 + 20Q$ .  
 Therefore, to maximize sales,  $R$ , we require

$$\frac{dR}{dQ} = -2Q + 20 = 0 \quad \text{or} \quad Q_r = 10,$$

and direct substitution yields

$$P_r = 10, \quad \Pi_r = -32, \quad R_r = 100$$

**Illustration 2:** Suppose that a firm has a linear demand function such as  $P = 20 - Q$  and total cost is  $TC = 0.5Q^2$ . The unit cost functions will be  $MC = Q$  and  $AC = 0.5Q$ . Find out the profit maximising level of price and output.

**Solution:**

The price/output combination that would maximize profit can be determined as

$$\pi = R - TC$$

$$\pi = P \cdot Q - TC$$

$$\pi = 20Q - Q^2 - .5Q^2$$

$$\frac{d\pi}{dQ} = 20 - 2Q - Q = 0$$

$$20 - 2Q = Q \quad (\text{i.e., } MR = MC)$$

$$3Q = 20$$

$$Q_{\pi} = 6.67$$

$$P_{\pi} = 13.33$$

$$\pi_{\pi} = 66.67$$

The price/output combination that maximizes sales revenue is found by simply taking the first derivative of revenue ( $R$ ) with respect to output ( $Q$ ).

$$R = P \cdot Q$$

$$R = 20Q - Q^2$$

$$\frac{dR}{dQ} = 20 - 2Q = 0$$

$$2Q = 20$$

$$Q_R = 10$$

$$P_R = 10$$

$$\pi_R = 50$$

### 3.3.4 The Dynamic Model:

The multi-period model of Baumol is based on the following assumptions:

- The objective of the firm is to maximize the rate of growth of sales revenue over its life cycle.
- There is no profit constraint; profit is the main source of financing growth of sales. Profit is thus an instrumental variable whose value is endogenously determined.
- Demand and cost curves have traditional shape; average revenue is downward-sloping and average cost is U-shaped.

Suppose sales revenue (R) grows at a rate of growth (g) per cent. Over its whole life the firm will have the following stream of revenues:

$$R, R(1 + g), R(1 + g)^2, \dots, R(1 + g)^n$$

The present value of this stream of future revenues can be computed by applying the usual discounting procedure.

$$R, R \left( \frac{1 + g}{1 + R} \right), R \left( \frac{1 + g}{1 + R} \right)^2, \dots, R \left( \frac{1 + g}{1 + R} \right)^n$$

where r is the rate of discount determined by the level of expectations and risk preferences of the firm.

The total present discounted value of all future revenues is expressed as:

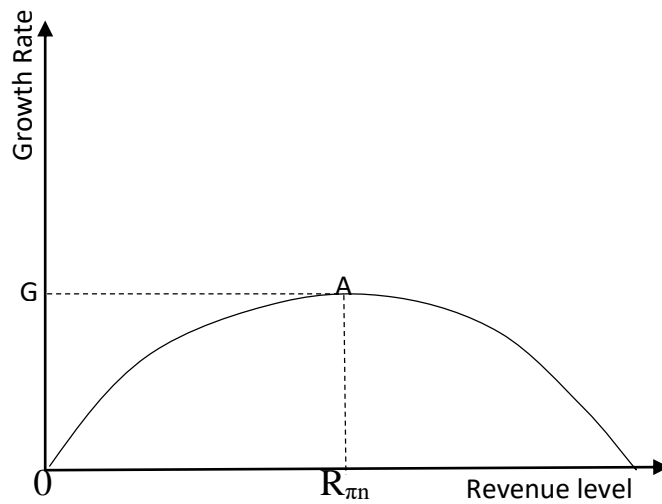
$$S = \sum_{i=0}^n R \left( \frac{1 + g}{1 + R} \right)^i$$

The firm seeks to maximize s by choosing an appropriate combination of current values of R and g. It is pretty obvious that

$$\frac{\partial S}{\partial R} > 0, \frac{\partial S}{\partial g} > 0$$

Also note that  $g = g(\pi, R)$  is the growth function and  $\pi = \pi(R, C, g, r)$  is the profit function. The growth function is derived from the profit function. Growth is mainly financed by ploughed back profits which depend on current level of revenue (R), cost(C), growth rate of sales (g) and the discount rate (r). To maximize S, the firm can choose a particular combination of R and g out of a set of alternatives.

These combinations are plotted along the growth curve, shown in figure 4. In this diagram up to point a, which corresponds to maximum profit level, R and g increase simultaneously. Beyond A, R increases but g tends to fall. Thus beyond  $R_{\pi}$  sales revenue level and growth rate become conflicting goals.



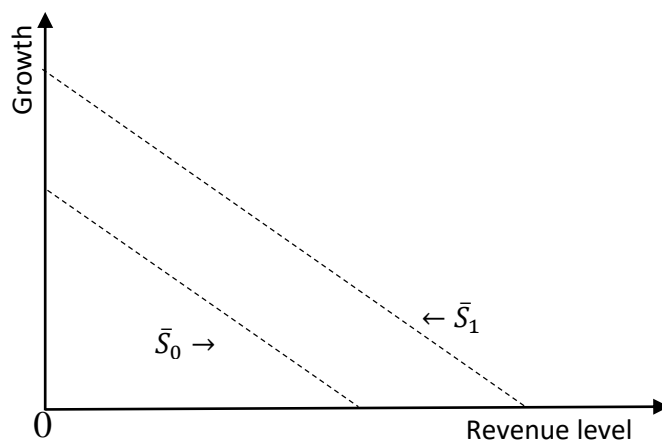
**Figure 3.19: The Growth Curve**

The optimum combination of R and g may not be a feasible one and vice-versa. Actual choice depends both on desirability and on feasibility. The desirability may be defined in terms of iso-present value curve. This curve is a locus of points showing alternative combinations of g and R which yield the same S.

Here S, the aggregate discounted present value of revenue, depends on R and g, given the exogenously determined discount rate. Thus we may assume that

$$S = a.g + b.R \dots \text{such that } g = \frac{1}{a}S - \frac{b}{a}R \text{ and } R = \frac{1}{b}S - \frac{a}{b}g$$

This is an equation of the iso-present value curve in the slope-intercept form. Thus, it is possible to think of a family of such curves, the highest one representing the maximum present value of S and the lowest one representing the minimum present value. The slope of this straight-line is given by a/b along a given curve, the level of S remains the same.



**Figure 3.20: The Iso-present value**

In order to choose the optimum combination of R and g, it is necessary to put the previous two diagrams together and design it as a case of growth-constrained iso-present value of revenue maximization.

In this case, the equilibrium solution is reached at point E at Figure 6 from which it appears that the firm will choose a combination of  $R^*$  and  $g^*$  to reach the highest possible level of  $S$ , subject to the growth function constraint.

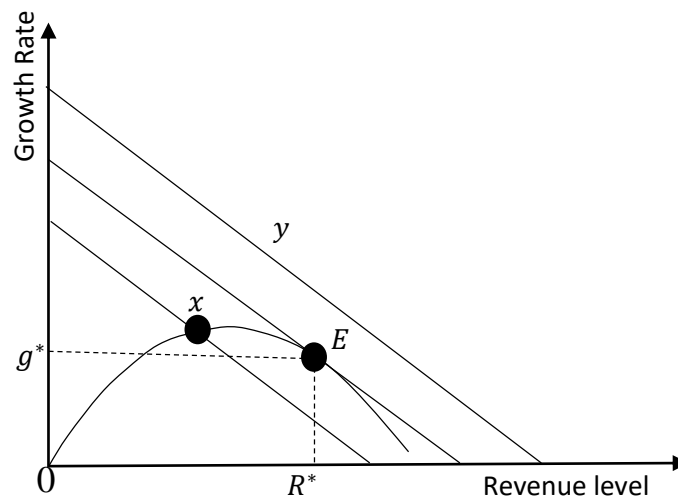


Figure 3.21: Growth- Constrained Revenue Maximization

### SELF-ASSESSMENT EXERCISE

Explain Baumol's Single Period Sales (Revenue) Maximization subject to Profit Constraint.

#### 3.4 Non-Optimizing Theories:

By criticizing the profit-maximization hypothesis modern economists have developed certain theories of the firm which do not hypothesize any optimizing behaviour. We have noted that the most celebrated managerial models are those of Baumol, Marris and Williamson.

They are distinguished primarily by the assumed objectives of the managers. Baumol suggested that managers maximise sales revenue, Marris, that they maximise growth, and Williamson, that they maximise a utility function including 'staff or 'emoluments'.

In each case the existence of monitoring from outside and limits to managerial discretion were explicitly recognised. Baumol included a minimum profit constraint in his model, and Marris similarly incorporated a valuation ratio constraint to reflect external pressure, i.e., from shareholders.

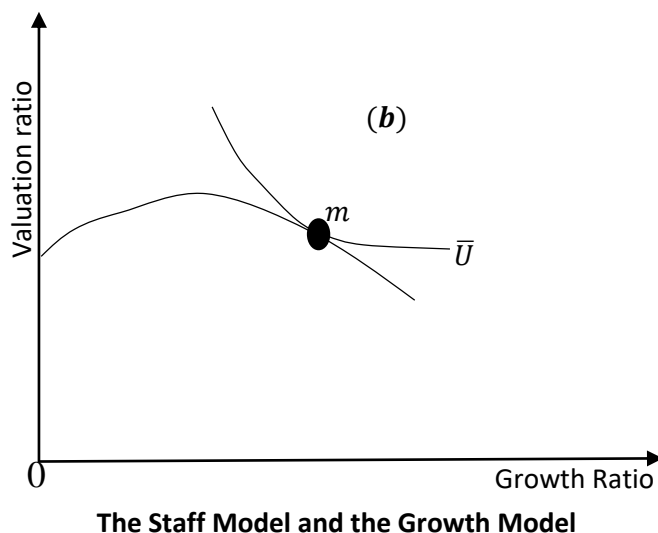
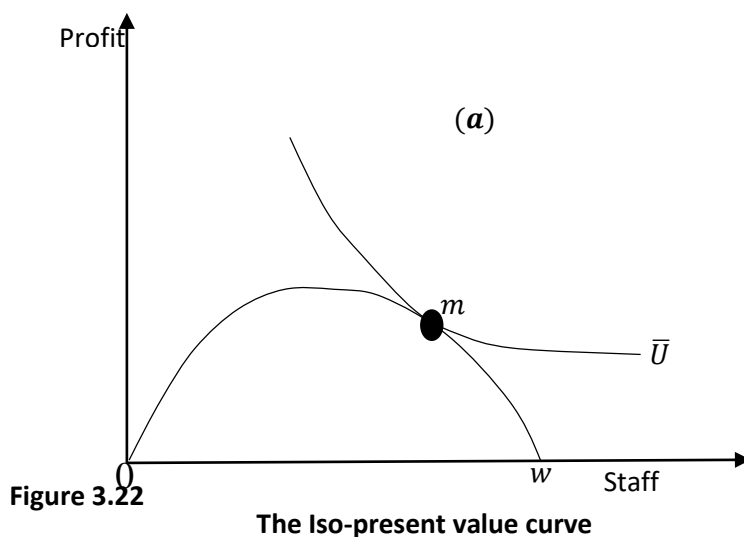
The valuation ratio is the market value of outstanding equity shares divided by the book value of the assets of a firm. According to Marris too low a ratio will involve a risk of takeover 'unacceptable' to the management.

In many ways figure 6 is absolutely typical of diagrammatic representations of managerial models of the firm. In Williamson's managerial firms the constraint OW [Figure 6 (a)] is derived as the summation of marginal revenue minus marginal cost.

In other words Williamson's firm is a monopolist. For Marris the diagram is again basically the same with the horizontal axis now measuring the rate of growth and the vertical axis the valuation ratio [Figure 6(b)].

The constraint is not supposed to emanate from the origin but is likely to have the same concave shape.

If growth is pushed past a certain point the value of shares on the market will fall as diseconomies associated with staff training are encountered (Penrose effects) and as a greater proportion of earnings are retained in the firm to finance expansion instead of being distributed as dividends to shareholders.



Being dissatisfied with the profit-maximization models of economists in 1955, H. A. Simon (the 1978 Nobel Laureate in Economics) has put forward the hypothesis that firms run by single enterprisers (who are also the owners) are likely to have different objectives from firms operated by modern executives in large corporations (who are supposedly not the owners).

Simon argues that managers in most cases have imperfect knowledge and inadequate information on the basis of which to take decisions.

In fact, if perfect knowledge and complete information were not available, the calculations involved in the decision-making process would be too complex to be practicable; and that, given this and the other inevitable uncertainties surrounding the decision making process in reality, business people can never be confident whether they are maximizing profits or not. Instead, business people “satisfice” rather than maximize, i.e., their aim is to earn just satisfactory profits.



Simon basically puts forward the proposition that firms have an 'aspiration level' which they seek to reach. In fact, what the satisfactory aspiration level of profits will be depends on past experience and will take account of future uncertainties, too. This level may be in terms of sales, market share, profits, etc. For any fixed period of time actual results are compared with the aspiration level.

If actual performance exceeds the aspiration level, no corrective action is called for. Instead the aspiration level for the next period is revised upward. On the contrary, if actual performance falls short of the aspiration level, the firm attempts to identify or search out the causes of discrepancy by spending sufficient time, effort and money.

Alternatively, if no apparent inefficiency is found (and the shortfall is believed to be due to external factors — factors beyond the control of the firm or its management) the firm will be constrained to revise its aspiration level for the next period downward. The aspiration level is, of course, the consensus of what can reasonably be expected in near future in the light of past performance.

However, since the cost of gathering information is high, all the alternatives will not be explored. A satisfactory alternative course of action is likely to be selected. This will probably not be the profit-maximising alternative.

Simon also argues that if neither search behaviour nor the lowering of aspiration levels quickly results in the achievement of a 'satisfactory' situation then the manager's behaviour pattern will become one of apathy or of aggression. In this sense this model does not have managerial usefulness.

Simon has also argued that the effort of trying to squeeze the last rupee of profitability out of the operation of the firm is likely to put extra strains and stresses on the business manager which in most cases, may not be liked by him.

He therefore seeks to reach a level of profit which yields an income which he regards as satisfactory and does not put any special effort to extract any extra naira of benefit. He satisfies rather than maximises.

The validity of Simon's hypothesis (i.e., the desire of businessmen for quiet life) largely depends on the business environment. In a highly competitive environment, a businessman has to work hard in order to safeguard his position (and thus protect his market share), whether he likes it or not.

On the contrary, if there is not much competitive rivalry in the area of business in which he is operating, he can afford the luxury of 'quiet life' and Simon's hypothesis may carry enormous good sense.

However, a related point may be noted in this context. In a single owner firm (i.e., sole proprietorship concern) it is possible for the owner-manager to 'satisfice' rather than maximize. But it is not possible for the head of a managerial team in a joint stock company to behave like this.

He may well be subject to various pressures from below to pursue a more expansionist policy. The pressure may come from those who are ambitious but are placed less comfortably than he is (i.e., at a lower point in the organization chart).

Shareholders may also demonstrate this type of ‘satisficing’ behaviour. A private shareholder is always at liberty to sell the share of a company if he is not satisfied with its performance and feels that he can secure a better return on his investment elsewhere.

But he is usually constrained by a lack of information. Thus he tends to act as a ‘satisfier’ so that if, for instance, dividends are held at a customary level, shareholders do not usually inquire whether they should be higher if management were better.

### **3.4.1 The Behavioural Theory of the Firm:**

In their book *A Behavioural Theory of the Firm* (1963), Cyert and March go a step ahead of Simon in making an in depth study of the way in which decisions are made in the large modern (multi- product) firm (characterized by divorce of ownership from management) under uncertainty in an imperfect, market.

They have expressed more concern in the decision making process than in the objectives or motivations of such firms (e.g., profit/sales maximization and satisficing). They look at the bureaucratic structure of the firm and study the nature of interrelationships of its various parts.

At the outset Cyert and March declare that if we are to develop a theory that predict and explain business decision making behaviour, the following two points have to take note of:

- (i) People (i.e., individuals) like organizations have goals,
- (ii) In order to define a theory of organizational decision making, we need something analogous — at the organizational level — to individual goals at the individual level.

Cyert and March set the formation of organizational goals through the notion of a coalition. The firm itself is visualized as a coalition of individuals who are organized in sub-coalitions. So they differ from Baumol and Simon who have assumed that the firm is dominated by a single person who makes the decisions and whose authority is unquestioned.

Instead Cyert and March assume that the firm is a decision making entity in its own right. They have recognized that management must achieve an objective, or possibly a set of objectives, through the efforts of a group of persons or through a coalition.

The coalition consists of the various units or parties associated with the firm such as managers, workers, shareholders, customers, suppliers as also professional people like accountants, auditors, lawyers, etc.

As with most others, such coalitions are not necessarily stable. Membership may change over time and also when particular decisions are involved. Within any group there is unlikely to be any permanent unanimity of purpose, although it may be worthwhile or expedient to act for a time as though there were. There is still less chance of acceptance of the goals of the firm by all the members of the coalition.

Thus the overriding problem of the leader of the coalition, who may be designated as the entrepreneur, is to attempt to resolve the conflict of goals and to keep all members pulling, more or less, in the same direction as long as possible. However, he must always be prepared for an unforeseen situation or sudden emergency.

The starting point of the behavioural theory is “where the entrepreneur makes a contract with the individual whereby the latter agrees to carry out instructions and to accept the organizational goal, or goals, as interpreted by the entrepreneur.” In order to get full support from the subordinates, the entrepreneur has to make ‘side payments’.

Alternatively put, the goals keep on changing through a process of bargaining, in which side payments are involved. Side payments not only involve money but nonpecuniary benefits also like authority, personal treatment, etc.

At the management level these involve matters outside the normal contract of employment (salaries, paid holidays, hours of work, etc.). The most important one seems to be policy commitments of one kind or another. This is known as policy side payment.

Finally, a winning coalition forms and the goals are set. However, the position is not static. Due to continuous changes in circumstances the bargaining is going on most of the time so that the coalition and its goals are liable to alter frequently.

In other words, a process of unrestricted bargaining would be inconsistent with stability in the organization. However, stability can be secured by working outside payments for those situations that are thought likely to occur.

There is, of course, likely to be conflict within such a coalition. Thus it is quite likely that some of the goals may be incompatible. However, such conflict resolution is possible in two ways. Firstly, decisions may be decentralised into divisions and departments.

Therefore conflict may be isolated geographically to ensure that all conflicts do not arise within the same unit. Secondly, crises and conflicts may be dealt with sequentially, i.e., they can be spaced out inter-temporally (i.e., over time) and can be tackled as and when they arise.

### **3.4.2 Goals of the Firm:**

Cyert and March go a step further and postulate that the firm may be pursuing the following five basic common goals:

**(a) Production goal:** This goal will be set as a target for the period and will have two aspects: level and smoothness. For example, a division may be set up to reach a specified goal (say, producing 100 units of a commodity per day) with the restriction that output should not deviate by more than 10% from this figure.

**(b) Inventory goal:** Business firms have to hold inventories because production and sales do not always coincide. It is absolutely essential to hold sufficient stocks of finished goods to meet consumer demand (as and when it arises). At the same time, it is to be ensured that there is no excessive stock holding at high cost.

This goal may be specified in terms of a target level and upper and lower limits may be set.

**(c) Sales goal:** This goal may be specified for the future either in volume or in value terms. Moreover it may again be expressed in terms of a level and/or range.

**(d) Market share goal:** The firm may set a target related to its share of the market (i.e., the industry of which it is a part for the product concerned). In some cases this may be a substitute for the sales goal, but in other cases it may be a supplementary goal.

**(e) Profit goal:** The purpose of setting this goal is twofold: to measure the effectiveness of management and to act as a source of payment of dividends to shareholders.

The behavioural theory does not postulate goal maximization but seeks sub-optimization or attainable goals. Like Simon, Cyert and March state that firms compare performance with goals. What will be sought at any time largely depends on the level of aspirations. If the goal is met no action will be taken. But in practice the level of aspirations, in most cases, outstrip achievements.

In contrast, if achievement improves rapidly, then it may outstrip aspirations, which may then be revised upwards. In a like manner, where achievement worsens there may be a tendency for a downward revision of aspirations to occur. There is thus likely to be a certain adjustment of goals in the light of experience.

If, however, performance falls short of aspirations (i.e., the goal is not met) a search activity is initiated to identify the causes of non-attainment. If the reason is within the firm's compass, steps are taken to rectify the non-attainment (i.e., alternative courses of action will be stimulated). This imposes extra costs on the firm and will not be carried beyond the point where a satisfactory solution is found.

If a number of alternatives are found, the best one will be selected and no additional search will be carried out to see whether any further improvement is possible. If the reason is outside the control of the firm (e.g., depressed market conditions due to recession in the economy) the goal for the next period may be revised downward.

### **3.4.3 Cyert and March Organizational Slack**

Cyert and March argue that the coalition will remain viable so long as the payments are sufficient to keep the members within the organization. So it is absolutely essential to develop a satisfactory 'package' of money together with other benefits which will prevent the individual manager from looking for openings elsewhere.

In practice, however, there is likely to be disparity between the actual payment which is made and that which is necessary to keep the individual in the organization.

However, it is not that easy to calculate side payments accurately. Usually payments made, tend to exceed what is really necessary. Such excess payment is termed organizational slack. The concept is of considerable importance in rectifying the non-attainment of goals.

The following three examples bear relevance in this context:

- (i) Shareholders are often paid more than what is required to keep them holding shares.
- (ii) Wages are often in excess of those required to keep workers within the organization.
- (iii) Executives in most cases are provided with luxuries and services in excess of what they really need.

Cyert and March argue that organizational slack (OS) grows naturally as the firm itself grows and prospers over time; it is not a deliberate objective. However, when circumstances become more and more adverse, OS provides the first means of making economies on costs.

Under difficult conditions there will be real pressure to reduce those side payments which can no longer be afforded at their original level. This slimming operation will, in all likelihood, reduce the organizational slack, while, at the same time, still leave the members of the organization sufficiently satisfied to stay within it.

#### **4.0 CONCLUSION**

From our discussion so far, we can infer that: The behavioural approach of Cyert and March is a dynamic one.

Three major points that emerge from the approach are as follows:

1. The goals and objectives of a firm will emerge from the coalition in existence, at any given point of time.
2. However, there is likely to be a change in coalition, and with it, the objectives pursued by the organization as a whole.
3. Hence, not only different firms will have different objectives at the same point of time, but the same firm may have different aims and objectives at various time periods.

#### **5.0 SUMMARY**

The theory of the firm is the microeconomic concept that states the overall nature of companies is to maximize profits meaning to create as much of a gap between revenue and costs.

The theory has been debated as to whether a company's goal is to maximize profits in the short-term or long-term.

Solely focusing on profit maximization comes with a level of risk in regards to public perception and a loss of goodwill between the company, consumers, investors, and the public.

#### **6.0 TUTOR-MARKED ASSIGNMENT**

1. The behavioural theory does not postulate goal maximization but seeks sub-optimization or attainable goals. Explain.
2. Why is profit maximised when  $MC = MR$ ? and what is the basic assumption that economists make about the objectives of private firms?
3. Suppose that a firm has a linear demand function such as  $P = 20 - Q$  and total cost is  $TC = 0.5Q^2$ . The unit cost functions will be  $MC = Q$  and  $AC = 0.5Q$  and the firm wants to maximize sales revenue subject to a specific rate of profit (perhaps designed to achieve a particular rate of return). We assume that the firm wishes to constrain their profit to 20 percent of the total unit cost of production. Find out the level of profit.

#### **7.0 REFERENCES/FURTHER READING**

Chris B. Murphy (2019), Theory of the Firm. ). Retrieved on 22/05/2020 from <https://www.investopedia.com/contributors/99607/>

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

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

Lipsey, et al. (2008). *Economics*. (13th ed.). Upper Saddle River, NJ: Pearson.  
Conclusion:

Samia Rekhi, Top 3 Theories of Firm (With Diagram). Retrieved on 22/05/2020 from <https://www.economicdiscussion.net/firm/top-3-theories-of-firm-with-diagram/19519/>

## **UNIT 2: THE THEORY OF DISTRIBUTION**

### **CONTENTS**

- 1.0 Introduction
- 2.0 Objectives
- 3.0 Main Content
  - 3.1 Concept of Theory Distribution
  - 3.2 Reason for the Study of Factor Pricing
  - 3.3 The Differences between Factor Market and Product Market
    - 3.3.1 Nature of Demand
    - 3.3.2 Nature of Supply in Relation to Price
    - 3.3.3 Nature of Supply in Relation to Cost
- 4.0 Conclusion
- 5.0 Summary
- 6.0 Tutor-Marked Assignment
- 7.0 References/Further Reading

### **1.0 INTRODUCTION**

The theory of distribution is that incomes are earned in the production of goods and services and that the value of the productive factor reflects its contribution to the total product. Distribution refers to the way total output, income, or wealth is distributed among individuals or among the factors of production such as labour, land, and capital. In general theory and the national income and product accounts, each unit of output corresponds to a unit of income. It is the systematic attempt to account for the sharing of the national income among the owners of the factors of production i.e., land, labour, and capital. Economists have studied how the costs of these factors i.e., rent, wages, and profits and the size of their return are fixed.

### **2.0 OBJECTIVES**

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

- ✓ Understand the concept of theory of distribution
- ✓ Explain problems associated with the theory of distribution
- ✓ Determine the various uses of marginal productivity.
- ✓ State assumptions of theory of marginal productivity.

### **3.0 MAIN CONTENT**

#### **3.1 Concept of Theory of Distribution**

The theory of distribution is an attempt to explain how income is distributed among the factors of production. To produce goods and services, firms need the services of factors of production, such as land, labour, capital, and entrepreneur. The prices paid for these factors of production are called rent, wage, interest, and profit.

In the modern time, the production of goods and services is a joint operation. All the different factors of production i.e., land, labour, capital and enterprise are combined together in productive activity.

Productive activity is thus the result of the joint effort of these four factors of production which work collectively to produce more wealth. These factors need to be paid or rewarded for their services for producing the wealth.

### **3.1.1 Personal Distribution and Functional Distribution:**

In economics, the term ‘distribution’ has two components:

- (i) Functional distribution,
- (ii) Personal distribution.

#### **1. Functional Distribution:**

Functional distribution refers to the distinct share of the national income received by the people, as agents of production per Unit of time, as a reward for the unique functions rendered by them through their productive services. These shares are commonly described as wages, rent, interest and profits in the aggregate production. It implies factor price determination of a class of factors. It has been called as “Macro” concept.

However, the theory of functional distribution, which attempts to explain the prices of land, labour, and capital, is a standard subject in economics. It sees the demand for land, labour, and capital as derived demand, stemming from the demand for final goods. Behind this lies the idea that a businessman demands inputs of land, labour, and capital because he needs them in the production of goods that he sells. The theory of distribution is thus related to the theory of production, one of the well-developed subjects of economics.

#### **2. Personal Distribution:**

Personal distribution on the other-hand, is a ‘Micro Concept’ which refers to the given amount of wealth and income received by individuals in society through their economics efforts, i.e., individual’s personal earnings of income through various sources.

The concept of equality and inequality of income distribution and social justice is basically concerned with the personal distribution of income. Taxation measures are designed to influence personal distribution of income and wealth in a community.

The theory of distribution deals with functional distribution and not with personal distribution of income. It seeks to explain the principles governing the determination of factor rewards like—rent, wages, interest and profits, i.e., how prices of the factors of production are set.

Personal distribution is primarily a matter of statistics and the conclusions that can be drawn from them. The inequality seems to be greatest in poor countries and diminishes somewhat in the course of economic development. Some authorities point to the natural inequality of human beings (differences in intelligence and ability), others to the effects of social institutions (including education); some emphasize economic factors such as scarcity; others invoke political concepts such as power, exploitation, or the structure of society.



### 3.1.2 Importance of Distribution

At present under the study of economics the study of 'Distribution' has occupied a very important place. The methods and systems of distribution has high effect on the economic life of the nation. Therefore, where the work of distribution is done with equity and justice the various channels of distribution are satisfied with its workings.

The satisfied workers increase their efficiency and they increase the quality and quantity of production. Contrary to this if the methods of distribution are improper and a particular class is being exploited then there will be dis-satisfaction feeling will crop up among people. Therefore, with the study of the distribution, it is clear that in the country with scientific system of production, equity and scientific way of distribution method is also very essential.

### 3.1.3 Main Problems of Distribution

Main problems of Distribution are as follows:

1. How much property be distributed?
2. Among what factors it should be distributed?
3. What should be the theory of distribution?

**1. How much Property be Distributed?** Distribution is made of national income. But national income is of two types:

- (i) Gross National Income and
- (ii) Net National Income.

In any country in one year whatever is earned in connection with goods and service, their value interns of money is called Gross National Income. But it should be remembered that Gross National Income is never distributed. Distribution is always done of Net National Income.

To earn total income one has to incur certain expenses. Therefore, in the total national income, after the deduction of the expenses whatever is left out that is known as net national income and the balance of the remaining money is distributed among the various factors of production.

Therefore, Net National Income = (Gross National Income) – (minus) Cost of raw-materials + Replacement cost of fixed and circulating capital + depreciation and repairs of fixed capital + Taxes and insurance charges

**2. Among What Factors to be Distributed?** National income is distributed among the various factors of production like—land, labour, capital and enterprise. From national income the rent of land, wages of labourers, interest on capital and risk part of money to entrepreneur will be deducted and the balance left will be net profit which will be distributed.

**3. What should be the Theory of Distribution?** Regarding the distribution of net national income the following two principles are being adopted.

They are as follows:

- (i) Marginal Productivity Theory of Distribution.
- (ii) Modern Theory of Distribution.

### 3.2 Marginal Productivity Theory of Distribution

Marginal productivity theory of distribution is the most celebrated theory of distribution. It is the neo-classical theory of distribution and is derived from Ricardo's "Marginal principle". J.B. Clark, Marshall and Hicks are the main pro-pounders of this theory. Initially, the theory was propounded as an explanation for the determination of wages (i.e., the reward for labour) but, later on, it was generalized as a theory of factor pricing for all the factors of production.

"The theory states that the price of a factor of production is governed by its marginal productivity. To support this hypothesis, it analyses the process of equilibrium pertaining to the employment of input of various factors by an individual firm under perfect competition. In a perfectly competitive factor market, a firm can buy any number of units of factors of production, at the prevailing market price. Now, the question is: given the price of a factor, how much of each factor will he employ."

According to this theory, an entrepreneur or a firm will employ a factor at a given price till its marginal productivity tends to be equal to its price. It thus follows that the reward (price) of a factor tends to be equal to its marginal productivity.

Marginal Productivity Theory of Distribution is the reward of a factor equals its marginal product. Marginal product, also known as marginal physical product, is the increment made to the total output by employing an additional unit of a factor, keeping all other factors constant. If the increase in the output is multiplied by the prevailing price of the product, the result is the marginal value product of that factor. But it is better to measure marginal product of a factor in terms of its marginal revenue product (MRP) which may be defined as the addition made to total revenue resulting from the employment of one more unit of a factor of production, other factors remaining unchanged.

In other words, by the marginal productivity of a factor of production we mean the addition made to total output by the employment of the marginal unit i.e., the unit which the employer thinks just worth-while employing. At the margin of employment, the payment made to the factor concerned is just equal to the value of the addition made to the total output on account of the employment of the additional unit of a factor.

For example: If the prevailing wage is less than the marginal productivity, then more labour will be employed. Competition among employers will raise the wage to the level of marginal productivity. If on the other-hand, the marginal productivity is less than the wage, the employers are losing and they will reduce their demand for labour. As a result, the wage rate will come down to the level of marginal productivity. In this way by competition, wage tends to equal the marginal productivity. This applies also to the other factors of production and their rewards.

Thus, it must be noted that in a position of competitive equilibrium:

- (a) The marginal productivity of a factor of production is the same in all employments,
- (b) The marginal productivity of a factor of production is measured by the price of the factor of production; and
- (c) Marginal productivities of various factors are proportional to their respective prices.

Further, over the whole field of employment, therefore, each factor of production tends to be paid in proportion to its marginal productivity. Thus, the distribution of national income or the total aggregate output of an economy is not a scramble as the strikes or lock-outs make it appear to be. It is governed by a definite economic principle viz. marginal productivity.

### **3.2.1 Assumptions of Marginal Productivity Theory**

The Marginal Productivity Theory of distribution is based on the following implicit and explicit assumptions:

- (i) There is perfect competition, both in the product market as well as in the factor market.
- (ii) There should not be any technological change. Therefore, the techniques of production should remain the same, though the scales and proportions of factors may change.
- (iii) All units of a factor should be perfectly homogeneous i.e., they should be of equal efficiency. This means that all units of a factor should receive the same price. The homogeneity of factors of units should imply that they are perfectly substitutes of each other.
- (iv) The firm aims at maximisation of profit. Therefore, it should seek and observe the most efficient allocation of resources.
- (v) The economy as a whole, should operate at the full employment level.
- (vi) There should be perfect mobility of factors of production.
- (vii) The bargaining power of the seller and the buyers of a factor of production should be equal.
- (viii) The marginal productivity of an individual should be measurable.
- (ix) There should not be any government intervention in the fixation of factor price, such as minimum wage legislation or price control etc.
- (x) The theory essentially considers long-run analysis in order to prove that the price of a factor will tend to be equal to both average and marginal productivity.

### **3.2.2 The Concepts of Productivity:**

Productivity means the quantity of the output turned out by the use of factor or factors of production. For example: How much wheat can be produced on 5 hectares of land under certain conditions or how much earth-digging can be done by 10 labourers.

Productivity of a factor may be viewed in two senses:

- (i) Physical productivity, and
- (ii) Revenue productivity.

#### **(i) Physical Productivity:**

Physical productivity of a factor is measured, in terms of physical units of output of a commodity produced by it per unit of time. When physical productivity is expressed in terms of money it is called revenue productivity. Again physical productivity has two concepts:

- (a) Average Physical product, and
- (b) Marginal Physical product.

(a) **Average Physical Product:** The average physical product or the average product of a factor is the total product divided by the number of units of the factor employed in the process of production. To put this in symbolic terms

$$AP = \frac{TP}{n}$$

(b) **Marginal Physical Product:** The marginal physical product of a factor is the increase in total product resulting from the employment of an additional unit of that factor, other factors remaining constant. The physical product or the marginal product of a particular factor is thus measured as  $MPP = TPP_n - TPP_{n-1}$  or  $\frac{\Delta TP}{\Delta L}$ .

MPP = Marginal physical productivity

TPP<sub>N</sub> = Total physical productivity of a factor and n is the units of the factor employed

TPP<sub>n-1</sub> = Total physical productivity of a factor and n-1 are the units of the factor employed

ΔTP = change in total product

ΔL = change in factor (labour)

Once the average and marginal products are calculated it is easy to measure the respective revenue productivity of the factor concerned. Here, we measure the quantity of the product in physical terms.

For example: We may express in terms of quintals of wheat or the number of chairs produced. But we are not concerned here with the total quantity of wheat or the average yield. We are concerned here with the marginal product which means an addition made to the total output of the commodity by the addition of one unit of a factor of production.

Suppose 3 hectares of land yield 30 quintals of wheat and 4 hectares, 40 quintals. The use of the third hectare has added 10 quintals. This is the marginal physical product. The total product has been increased by 10 quintals by the employment of the third or the marginal hectare. That is why it is called marginal product. But it is the physical product and not product in terms of value.

### 3.2.3 Value of Marginal Product (VMP)

This is also called Value of Marginal Physical Product (VMPP) and is usually referred to as the marginal productivity of a factor, and is obtained by multiplying the marginal physical product of the factor by the price of output.

Marginal Productivity of VMPP = MPP x P where, MPP stands for the marginal physical product of the factor, and P for the price of output. The marginal value product means the value of additional product obtained by the employment of another unit of a factor of production. We can get value product by multiplying the physical product i.e., the quantity of the commodity by its price in the market.

For example: When we say that it is an addition to the total product by the addition of one more unit of a factor of production, say one hectare or one worker, or a unit of ₹1,000 in capital. When this marginal product is expressed not in physical terms but in terms of its value in the market, it is called Marginal Value Product.

### 3.2.4 Marginal Revenue Productivity

The marginal revenue at any level of firm's output is the net revenue earned by selling another (additional) unit of the product. Algebraically, it is the addition to total revenue earned by selling  $n$  units of product. In other-words, Marginal Revenue Product (MRP) of a factor is the net addition to total revenue made by the employment of an additional unit of that factor, assuming other factors to be fixed under a given state of technology. Thus, marginal revenue product is obtained by multiplying the marginal revenue.

$$MRP = MPP \times MR$$

where, MRP indicates marginal revenue product, MPP stands for the marginal physical product and MR stands for the marginal revenue.

Thus, there is a conceptual difference between marginal revenue product (MRP) and value of marginal physical product (VMPP). In the former, we consider marginal revenue to be multiplied by the MPP and, in latter, we take price to multiply it by the MPP.

In perfectly competitive market conditions for the product, however,  $MPP = VPP$ . This is because under Perfect Competition—Price = MR. But if the commodity-market has imperfect competition price or AR tends to be greater than the marginal revenue, then VMPP will be higher than MRP.

### SELF-ASSESSMENT EXERCISE

From the above self-assessment exercise, if 1000 shirts produced by 100 labourers (with 10 units of capital) earned a revenue of N100000 (each was sold at the rate of N100), and if 1100 shirts were produced by 101 labourer earned the producer a revenue of N11000, calculate the MRP.

### 3.2.5 Criticisms of the Marginal Productivity Theory

Most of the economists are of this opinion that though the marginal productivity theory is logically sound and perfect, it has many inherent shortcomings and they have criticised the theory on the following grounds:

1. The Basic Assumption Underlying the Theory is Unrealistic: The theory is based on the assumption of perfect competition in the product as well as factor markets. Modern economists, like—Mrs. Robinson and Chamberlin have rightly pointed out that perfect competition is not a very large relative phenomenon. In reality, there is imperfect competition in the market. Further, other assumptions of the theory have also been criticised and they are as such.
2. All Units of Factor are not Homogeneous: The theory assumes that all units of a factor are homogeneous. In reality, however, all factor units can never be alike. Especially, the different labour units differ in efficiency and skill. Similarly, plots of land differ in fertility and so on.

3. Factors are not fully Employed: The theory assumes that all factors are fully employed. But, as Keynes pointed out, in reality there is a likelihood of under-employment rather than full employment.
4. Factors are not Perfectly Mobile: Next, the theory assumes perfect mobility of factors. But in reality, factors are imperfectly mobile between regions and occupations. There is no automatic movement of factors units from one place to another. The greater the degree of specialisation in an industry, the less is the factor mobility from one industry to another.
5. All Factors are not Divisible: The theory assumes the divisibility of factors. But lumpy factors like factory plant, machines and the manager are indivisible. In a large factory the addition or sub-traction of one factor units will have practically no effect on the total productivity. It may be true in domestic production. Thus, the equality between marginal productivity and price of a factor cannot be brought about by varying its quantities a little less or more.
6. This Theory is not Applicable in the Short-run: The theory is applicable only in the long-run, when the reward of a factor service tends to equal its marginal revenue product. But in reality, we are concerned with short-run problems. As said by Prof. Keynes— “In the long-run we are all dead.” This assumption makes the problem of pricing the factor-services unrealistic.
7. This Theory is a Static Theory: The marginal productivity theory is applicable only to a static economy as it regards no change in technology. Since the modern economy is dynamic and there are technological advances from time to time, the theory becomes in-applicable to modern conditions.
8. This Theory has been considered as One-sided: Because it considers only demand for factors in terms of its Marginal Revenue Product but it fails to analyse the conditions of supply in the factor market. The factor price may be high when the factor is relatively scarce.
9. Marginal Productivity of all Factors cannot be Measured Separately: In this theory it has been assumed that the marginal physical product of an individual factor can be measured by keeping other factors unchanged. Critics have said that one cannot consider the specific marginal productivity of a factor in isolation, when production is not the result of only one factor. It is the outcome of collective efforts of all factors at a time. Therefore, it is difficult to measure the marginal productivity of each factor separately. Since variation in output cannot be attributed to a single factor alone, marginal productivity appears to be a make-believe concept.
10. The Theory is based on the Law of Diminishing Returns as Applied to the Organisation of a Business: This means that a factor like capital with improved technology has increasing returns and it also enhances the productivity of other factors like labour. This theory misses this vital point of practical consideration.
11. Wage Determination Theory: This Theory has been criticised by Keynes and he is of this Opinion that theory is Basically Explained for Wage Determination and is

Loosely Extended for Pricing of the Other Factors of Production. But other factors like rent and capital have their distinctive factors like—rent and capital have their distinctive characteristics, so their rewards are also fixed distinctly. Again, the entrepreneur earns profit which is a residual income, which can be negative as well. Then, is it not ridiculous to lack of negative marginal product of an entrepreneur to explain loss in the business, which is improper.

12. The Theory cannot apply to Personal Distribution: The theory only explains functional distribution. It does not deal or explains anything of personal distribution of income and inequalities of earnings.

13. The Theory Lacks Normative Aspect of the Dealings: This theory contains only the positive aspect of the analysis. It does not consider anything or it does not have any ethical justification or social norm in determining the reward factor.

### **SELF-ASSESSMENT EXERCISE**

Describe two assumptions of marginal productivity theory.

### **3.3 Modern Theory of Distribution Demand and Supply Theory:**

We have seen earlier that the marginal productivity theory only tells us that how many workers an employer will engage at a given level in order to earn maximum of profit. It does not tell us how that wage-level is determined. Further, the marginal productivity theory describes the problem of the determination of the reward of a factor of production from the side of demand only. It has not said anything from the supply side.

Therefore, the marginal productivity theory cannot be said to be an adequate explanation of the determination of the factor prices. The modern theory of pricing which gives us a satisfactory explanation of factor prices in the Demand and Supply Theory. As we are aware that the price of a commodity is determined by the demand for and supply of, a commodity, similarly the price of a productive service also is determined by demand for and supply of that particular factor.

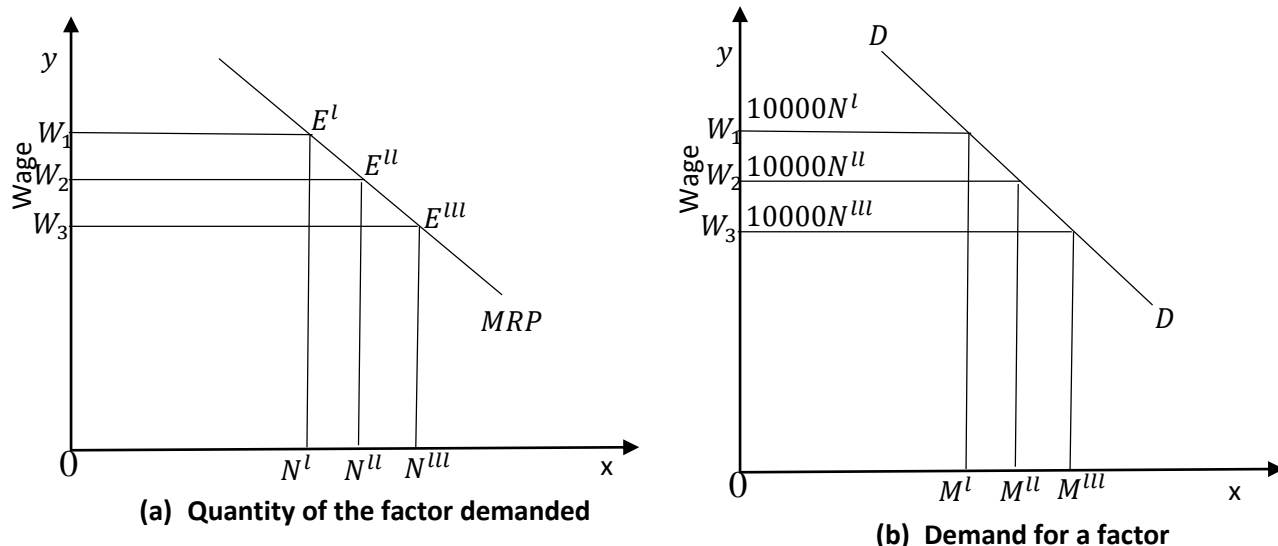
#### **3.3.1 Demand for a Factor:**

First we are going to consider the demand side of the factor. Here, we should remember that the demand for a factor of production is not a direct demand. It is an indirect or derived demand, it is derived from the demand for the product that, the factor produces. For example, we can say that labour does not satisfy our wants directly. The demand for labour entirely depends upon the demand for goods. If the demand for goods increases, the demand for the factors which help to produce those goods will also increase.

The demand for a factor of production will also depend on the quantity of the other factors required for the process. The demand price for a given quantity of a factor of production will be higher, the greater the quantities of the co-operating productive services. If in production more of a factor of production is employed, the marginal productivity of the factor will fall and the demand price will be lower of the unit of a productive service.

Further, the demand price of a factor of production also depends upon the value of the finished product in the production of which the factor is used. The demand price of a commodity is normally higher, if more valuable is the finished product in which the factor is used. Next, the more productive the factor is, the higher will be the demand price of a given quantity of the factor.

From the following diagram the given explanation given can be explained:



In the diagram given the wage is  $OW$ , the firm is in equilibrium at the point  $E$  and the demand for the factor is  $ON$ . Similarly at  $OW'$  wage the demand is  $ON'$  and at  $OW''$  the demand is  $ON''$ .  $MRP$  (Marginal Revenue Productivity) curve is the demand curve for a factor of production by an individual firm.

For determining the price of a factor, it is not the demand of the individual firm that matters but it is the total demand, i.e., the sum-total of the demands of all firms in the industry. The total demand curve is derived by the total summation of the marginal revenue productivity curves all the firms. This curve  $DD$  is shown in the figure. Thus, from this figure it can be ascertained that-according to the law of diminishing marginal productivity, the more a factor is employed, the lower is the marginal productivity.

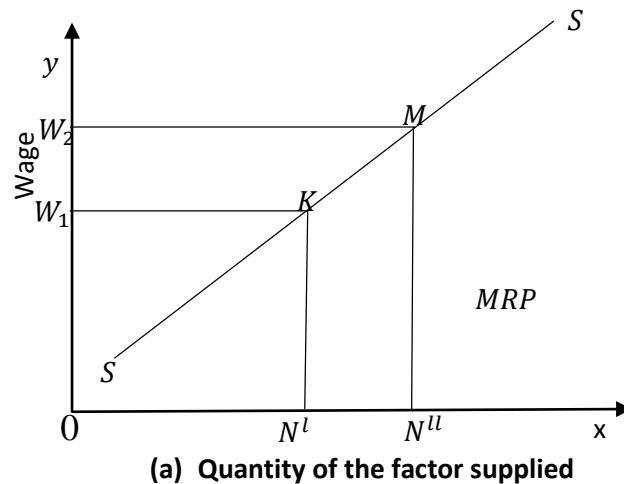
### 3.3.2 Supply Side

The supply curve of a factor depends on the various conditions of its supply.

The supply of factors of production is very complicated because each factor presents a peculiar problem of its own. Land, for instance, is fixed in quantity and its total supply cannot be increased even if its price rises. However, for a particular use, its supply can be varied. Similar is the case with labour. The total supply of labour in the country depends upon various factors, such as size of population, labour efficiency, expenses of training and education, geographical distribution, attitude towards work, etc. The total supply of labour in the country is fixed but for a particular occupation it can be increased by drawing workers from other occupations and by increasing the working hours of the labour already employed. The supply of capital is also complicated as it depends upon the power and willingness of the people to save. The marginal efficiency of capital and the rate of interest also play a very important role in the supply of capital in the country.



In a nutshell, we can say that the supply of a factor is also a function of price. The higher the price of a factor of production, other things remaining the same, and the greater will be its supply and vice versa. The supply curve of a factor of production is positively inclined, i.e., its slopes upward from left to right as is shown below:



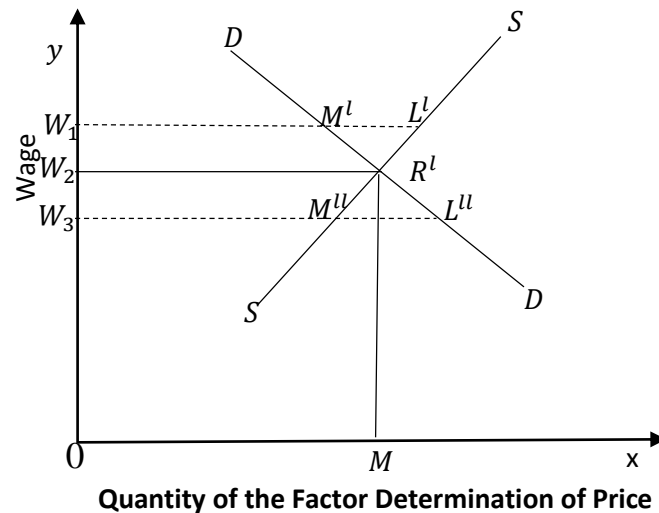
In the diagram above, we measure units of a factor, say labour, along OX-axis and wage on OY-axis. If the wage is  $OW_1$ ,  $ON^1$  workers are supplied. At wage  $OW_2$ , the supply of workers increases from  $ON^1$  to  $ON^{II}$ . The normal supply curve of a factor is positively sloped. It rises from left to-right upward indicating that at higher factor prices, greater quantity of factor is offered in the factor market and vice versa

For example: The supply of labour entirely depends upon the size and composition of population, the occupational and geographical distribution, labour efficiency their training, expected income, relative preference for work and leisure etc. By considering all these relevant factors, it is possible to construct the supply curve of a productive service.

Further, the supply of labour does not depend only on economic factors but many non-economic considerations also. Therefore, we can say that if the price of a factor increases, its supply will also increase and vice-versa. Hence, the supply curve of a factor rises from left to right upwards.

### 3.3.3 Factor Determination of Price

In a perfect competitive market, there are large number of firms to demand the services of a factor of production and also large number of households to supply the services of a factor. In such a factor market, the price of a factor is determined by the interaction of the forces of demand and supply as is shown in the figure below.



### Interaction of Demand and Supply

We have studied up to this stage the demand curve and the supply curve of the factor of production while in price fixation both curves are needed. Therefore, the price will tend to prevail in the market at which the demand and supply are in equilibrium. This equilibrium is at the point of intersection of the demand and supply curves.

In the diagram above the demand and supply curves intersect at the point R and the price of the factor will be OW at OW' demand W' M' is less than the supply W L'. In this case competition among the sellers of the service will tend to bring down the price to OW. On the other hand, at OW "price the demand W "L" is greater than the supply W "M ", hence price will tend to go up to OW at which the demand and supply will be equal.

To conclude, this is how that the price of a factor of production in the factor market is determined by the interaction of the forces of demand and supply in connection with the factor of production. Thus eminent economists are of this opinion that this is the proper, correct and satisfactory theory of distribution.

### SELF-ASSESSMENT EXERCISE

With the use of graphical illustration explain how the forces of demand and supply determine the price of factor in the factor market.

### 4.0 CONCLUSION

In many traditional economies, community interests take precedence over the individual. Individuals may be expected to combine their efforts and share equally in the proceeds of labour. In other traditional economies, some sort of private property is respected, but it is restrained by a strong set of obligations that individuals owe to their community.

Marginal productivity theory explains how factor prices are determined in markets, not how factor prices should be determined. However, marginal productivity theory has been criticised even as a positive theory. Since, product markets are not perfect, the marginal productivity theory may not always offer a realistic explanation of short-run factor pricing. However, over a long period of time, marginal productivity theory does offer a good explanation of factor prices.

Marginal productivity theory explains how factor prices are determined in markets, not how factor prices should be determined. However, marginal productivity theory has been criticised even as a positive theory. Since, product markets are not perfect, the marginal productivity theory may not always offer a realistic explanation of short-run factor pricing. However, over a long period of time, marginal productivity theory does offer a good explanation of factor prices.

## **5.0 SUMMARY**

The summary of the marginal productivity theory may thus be laid down in terms of the following propositions:

“The marginal productivity of a factor determines its price. In the long-run, the price or reward of a factor tends to be equal to its marginal as well as average products. When the reward of each factor in the economy tends to be equal to its marginal productivity, there is optimum allocation of resources (factors) in different uses. Further, when all factors receive their shares according to their respective marginal products, the total product will be exhausted.” In this unit, you learnt about marginal productivity theory, which states that a factor of production is paid price equal to its marginal product. You were also taught the assumptions of marginal productivity theory such as prevalence of perfect competition in factor as well as product market, all factors are identical, perfect substitute for each other, perfectly mobile, perfect divisibility of factors and so on. You also learnt that the demand for factor is a derived demand and supply for factor is directly related to price of offer. The equilibrium in this market is determined by the forces of demand and supply. You also learnt about factors that determine both the demand and supply of factor in the factor market.

## **6.0 TUTOR-MARKED ASSIGNMENT**

- i. Define marginal productivity of factor and explain its assumptions as well as its limitations. Assuming 1000 shirts were produced in a month when 100 labourers are engaged with 15 units of capital. When 101 labourers are employed with the same amount capital, 1100 shirts were produced. Calculate the marginal physical productivity (MPP).
- ii. Explain how the interplay of demand and supply in the factor market determine its equilibrium point.

## **7.0 REFERENCES /FURTHER READING**

Jain, T.R. & Khana, O.P. (2009). Business Economics. New Delhi, India:V.K. Publications.

Jain T.R. & Khana, O.P. (2009). *Managerial Economics*. New Delhi, India:V.K. Publications.

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

Robert, S. P. & Daniel, L. R. (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.

Saqib Shaikh, Distribution and Theories of Distribution (With Diagram)  
<https://www.economicdiscussion.net/theories-of-distribution/distribution-and-theories-of-distribution-with-diagram/13897/>

## **UNIT 3: COST-BENEFIT ANALYSIS**

### **CONTENTS**

- 1.0 Introduction
- 2.0 Objectives
- 3.0 Main content
  - 3.1 Meaning of Cost Benefit Analysis
  - 3.2 History of Cost Benefit Analysis
  - 3.3 Steps in Cost Benefit Analysis
  - 3.4 Cost incurred in Cost Benefit Analysis
  - 3.5 Cost and Benefits in Controlling Pollution
  - 3.6 Total Benefit Curves and Marginal Benefit Curves
  - 3.7 Optimum level of Environmental Quality
  - 3.8 Cost Benefit Analysis: The Framework
    - 3.8.1 Property Price Approach
  - 3.9 Merits and Demerits of Cost Benefit Analysis
- 4.0 Conclusion
- 5.0 Summary
- 6.0 Tutor-Marked Assignment
- 7.0 References/Further Readings

### **1.0. INTRODUCTION**

Cost-Benefit Analysis (CBA) estimates and totals up the equivalent money value of the benefits and costs to the community of projects to establish whether they are worthwhile. These projects may be dams and highways or can be training programs and health care systems.

Cost benefit analysis (CBA), sometimes called benefit costs analysis (BCA), is a systematic approach to estimating the strengths and weaknesses of alternatives (for example in transactions, activities, functional business requirements or projects investments); it is used to determine options that provide the best approach to achieve benefits while preserving savings. The CBA is also defined as a systematic process for calculating and comparing benefits and costs of a decision, policy (with particular regard to government policy) or (in general) project.

Broadly, CBA has two main purposes; to determine if an investment/decision is sound (justification/feasibility) – verifying whether its benefits outweigh the costs, and by how much; to provide a basis for comparing projects – which involves comparing the total expected cost of each option against its total expected benefits.

CBA is related to (but distinct from) cost-effectiveness analysis. In CBA, benefits and costs are expressed in monetary terms, and are adjusted for the time value of money, so that all flows of benefits and flows of project costs over time (which tend to occur at different points in time) are expressed on a common basis in terms of their net present.

### **2.0. OBJECTIVES**

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

- ✓ Define and understand the meaning of cost benefit analysis
- ✓ Understand the steps/cost incurred in cost benefit analysis

- ✓ Understand the cost and benefit in controlling pollution
- ✓ Understand the total benefit curves and marginal benefit curves.
- ✓ Understand the optimum level of environmental quality
- ✓ Understand the property price approach

### **3.0. MAIN CONTENT**

#### **3.1. Meaning of Cost Benefit Analysis**

The foundation of the method of cost benefit analysis arose from the Hicks – Kaldor criterion of efficiency maximization in 1939. The criterion of Hicks-Kaldor states that a project or activity merits consideration or remains desirable when the total benefits exceed total cost.

The cost benefit analysis states that if the benefit arrived from the pollution elimination programme is greater than the benefit received from it, it is called “positive benefit”. On the other hand, if the cost incurred in that programme is greater than the benefit received from it, it is termed as “negative benefit”.

The cost benefit analysis is the tool generally undertaken by the government for the welfare of the entire society. So it is also referred as social benefits. Cost benefit analysis may be summed as “the cost benefit analysis which involves measuring, adding up and comparing all the benefits and all the cost of a particular public project or a programme.”

#### **3.2. History of Cost-Benefit Analysis**

CBA has its origins in the water development projects of the U.S. Army Corps of Engineers. The Corps of Engineers had its origins in the French engineers hired by George Washington in the American Revolution. For years the only school of engineering in the United States was the Military Academy at West Point, New York.

In 1879, Congress created the Mississippi River Commission to "prevent destructive floods." The Commission included civilians but the president had to be an Army engineer and the Corps of Engineers always had veto power over any decision by the Commission.

In 1936 Congress passed the Flood Control Act which contained the wording, "the Federal Government should improve or participate in the improvement of navigable waters or their tributaries, including watersheds thereof, for flood-control purposes if the benefits to whomsoever they may accrue are in excess of the estimated costs." The phrase if the benefits to whomsoever they may accrue are in excess of the estimated costs established cost-benefit analysis. Initially the Corps of Engineers developed ad hoc methods for estimating benefits and costs. It wasn't until the 1950s that academic economists discovered that the Corps had developed a system for the economic analysis of public investments. Economists have influenced and improved the Corps' methods since then and cost-benefit analysis has been adapted to most areas of public decision-making.

### 3.3. Steps in Cost Benefit Analysis

Before entering into the cost benefit analysis we shall see the various steps taken in this analysis. The different Steps in Cost Benefit Analysis are:

- (i) Specify clearly the project or programme.
- (ii) Describe quantitatively the inputs and outputs of the programme.
- (iii) Estimate the social cost and benefits of these inputs and outputs.
- (iv) Compare these benefits and costs.

(i) **Specify Clearly the Project or Programme:** The first step is to decide on the perspective from which the study is to be done. Cost benefits are actually concerned with the public. When we have decided on the perspective, the main elements of the projects such as, the study of the location, timing, and groups involved, the connection with other programmes etc. should be considered. Again the pollution control phenomenon is a worldwide concept. So the regional planning agencies should give particular stress on the area of the study.

When the project is fixed the following two programmes are involved:

(a) **Physical project:** This implies the projects like the public waste treatment plants, beach restoration projects, hazardous waste removal, habitat improvement projects, land purchase for preservation etc. These projects are physical in nature, which is done when an area is polluted.

(b) **Regulatory projects:** This implies the enforcement of environmental laws and regulations, such as pollution standards, technical choices, waste disposal practice, restrictions of land for certain activities etc. This project regulates the amount of pollution in the society.

(ii) **Describe Quantitatively the Inputs and Outputs of the Programme:** The second step in cost benefit analysis is to determine the relevant force of input and output. For some projects it is easy to identify the input and output, For example, if we are planning a waste water treatment project, the staffs of that program will be able to provide a full physical specification of the plant, together with the inputs required to build it and keep it running.

However, it is harder to predict the externalities caused by the disposition of nuclear waste. A tolerable accuracy should be predicted with these projects. Because a restriction on development in a particular area can be expected to detect development elsewhere into the surrounding areas, since environmental projects or programmes don't usually last for a single year but are spread over for a long period of time.

(iii) **Estimate the Social Cost and Benefit of these Inputs and Outputs:** The next step is to put values on input and output flows i.e., to measure costs and benefits. We could do this in any units we wish but normally we take into account the monetary terms.

This does not mean in market value terms because in many cases we will be dealing with effects, especially on the benefit side that are not directly registered on market not it implies that only monetary values count in some fundamental manner.

This means that, we try to translate all the impact: of the project or the programme in order to make them comparable among themselves as well as with other types of public activities. Sometimes the projects or the programmes are immeasurable because we don't know how much value these projects or programme have in the economy.

**(iv) Compare these Benefits and Costs:** Next step is to make comparison between cost incurred and benefit derived from the project. One of them is to subtract the total cost from the total benefit to get net benefit. If the net benefit is positive, then the cost benefit is positive and if the net benefit is negative then the cost benefit is negative.

Again there is the other criteria called the cost benefit ratio. It is calculated by taking the ratio of benefit and cost. If the value is positive the cost benefit is positive and vice versa. These are certain steps to be taken into consideration before entering into the cost benefit project or programme.

### 3.4. Cost incurred in Cost Benefit Analysis

**(a) Pollution prevention cost:** It implies that the cost which is spent by the government or individuals or local bodies or firms in order to prevent pollution fully or partially. These types of costs are spent before the pollution is created in the society. For e.g., if in a particular area an industry is to be started then making arrangements so that the smoke or the waste water is disposed in an efficient way so that the waste product need not hit the society. Pollution prevention cost may be incurred whether in public or private sector. These costs are also called as Abatement Costs. The benefits arising out of these costs are called abatement benefit costs.

**(b) Pollution damage cost:** These are the costs incurred in the elimination of pollution that has already occurred for e.g., due to the industrial waste if a river nearby is affected then the government or private sector spend some money to clear that river. The amount is termed as the pollution damage cost.

**(c) Welfare damage cost:** If for e.g., an area is being polluted and the government or the private sector did not take any step to eradicate that pollution then it will lead to a damage in the welfare of the society. Therefore, pollution that is not prevented results in damaging the welfare of the society. So pollution that is not prevented results in welfare damage which may be pecuniary or real.

Summing all these costs the waste disposal cost is the sum of pollution prevention cost and pollution cost.

Pollution cost = pollution avoidance cost + welfare damage cost.

### 3.5. Costs and Benefits in Controlling Pollution

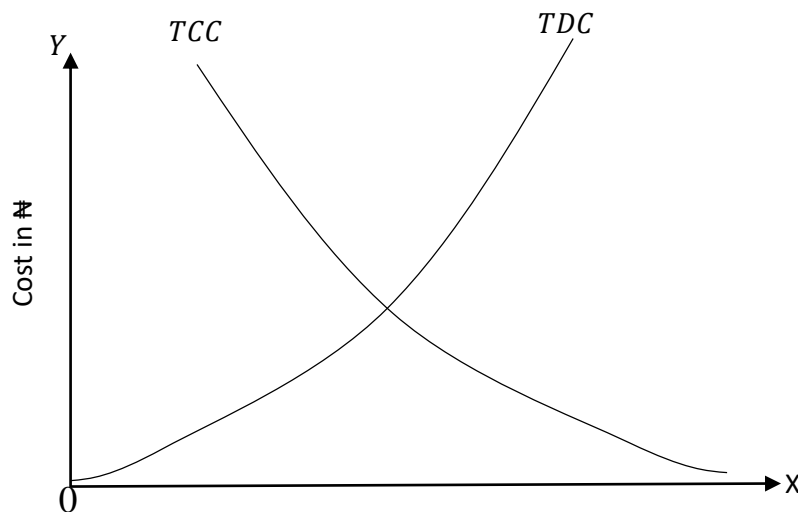
Pollution costs are mainly opportunity costs or real costs. Because if the pollution has not occurred this amount could be spent for alternative activity which gives welfare to the society. Actually these costs are resources utilized by reducing production of some other goods.

Figure 1. Illustrates the cost to the society due to the pollution. In the X axis we have the level of pollution and in the T axis are the cost involved for the elimination of pollution. If we go rightwards in the X axis it means the pollution increases and towards



left means less pollution. At the origin “O” the pollution is nil. Similarly, in the T axis if we go up, the cost incurred is high and lower the cost is less.

Figure 1: showing the cost to the society due to Pollution



**Figure 3.23: Amount of Pollution**

**Source: Adewole, (2016). Introduction to Public Sector Economics**

Keeping this we are drawing the total damage cost curve (TDC) which refers to the total external cost. This curve increases in an increasing rate stating that with the increase in the pollution the cost increases. The other curve is the total cost of pollution (TCC). This curve slopes upward from right to left. This implies that for acquiring less pollution we have to spend more money.

Having drawn the two curves, the good approach is to minimize the sum of TDC and TCC. In other words, by minimizing the sum of TDC and TCC we can acquire the optimum pollution and the social benefit will be high. This can be explained by the following method.

Let T be the output secured in the society with pollution control and  $Y_i$  be the flow without pollution control. The difference will be pollution cost, as pollution control incurs some cost which otherwise will be utilized for some production.

$$\text{So we say: } Y = Y_i - TCC$$

In the same way we can value the environmental quality service. This will be  $S_i$  without any pollution and S with such damages. The difference will be damage due to pollution.

$$\text{So we say: } S = S_i - TDC$$

Keeping this in mind the total social benefits are made up of the product produced in the country and the country’s environmental quality service. So,

$$\begin{aligned} \text{Total social benefit} &= Y + S \\ &= (Y_i - TCC) + (S_i - TDC) \\ &= (Y_i + S_i) - (TCC + TDC) \end{aligned}$$

$$\text{Therefore, } TSB = (Y_i + S_i) - (TCC + TDC)$$

In the above expression pollution affects TCC and TDC. Hence minimize the  $TDC + TCC$ . The figure coming below Figure 3.24 shows the optimum level of pollution by summing up the two cost curves and locating its minimum.

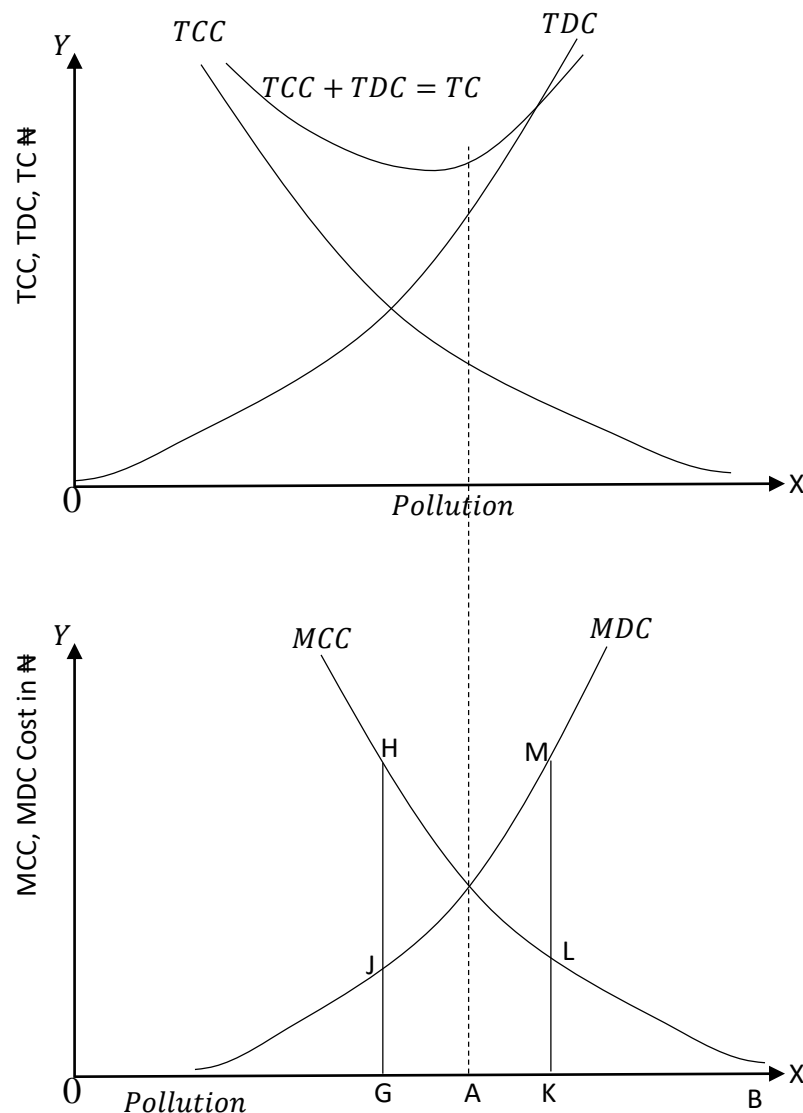


Figure 3.24: Optimum Level of Pollution

**Source: Source: Adewole, (2016). Introduction to Public Sector Economics**

Figure 3.24: showing the optimum level of pollution by summing up the two cost curves and locating its minimum.

In the figure the minimum point of TC curve is not the intersecting of total damage curve and total control curve. But it is at the point where marginal control cost curve and the marginal cost are in absolute magnitude. The optimum pollution is at A. Further reduction in pollution will cost more than its worth, when the pollution level exceeds A, extra cost to society of additional pollution is greater than the cost of preventing it.

The cost of allowing pollution to increase from A to B is much greater than the cost of preventing it. For instance, if pollution level is ok, managerial cost of controlling it is KL and marginal damage to society is KM and the KM is greater than KL.

To the other side cost of pollution control is greater than the cost to society of pollution. At G marginal cost of controlling pollution is GH and marginal damage cost of pollution

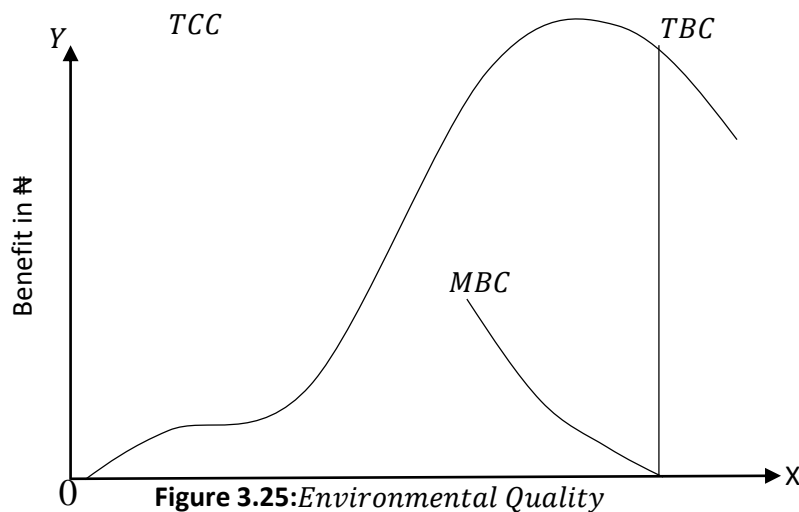
to society is GJ and GH is greater than GJ. So only at A the total costs are minimum and A is the optimal level of pollution.

### 3.6. Total Benefit Curves and Marginal Benefit Curves

When the environment is polluted highly the cost incurred in correcting is very high, so the TBC increases sharply. As the time passes the total benefit increases slowly as the marginal benefit arrived from it declines.

At last it reaches the minimum point and then shows a dropping tendency. So the MBC is a descending curve, where TBC is maximum the MBC cuts the X-axis. The total benefit increases at a decreasing rate. This can be seen in the diagram Fig. 3.

Figure 3: showing the total benefit curves and marginal benefit curves



Source: Wash, (2017). *Principles and Practice of Public Sector*

### 3.7. Optimum Level of Environmental Quality

The optimum level of environmental quality can be obtained when the two conditions are satisfied, they are:

- (a) The total benefit must be greater than the total cost.
- (b) The marginal benefit curve must be equal to the marginal control cost curve.

Keeping these conditions in mind we can draw a figure to explain the efficient level of environmental quality. In the figure 4. X-axis denotes the environmental quality and Y-axis represent the cost/benefit. The marginal control cost is rising.

The MCC and TCC indicate the opportunity cost of controlling pollution and the increase in the environmental quality. For the first condition, i.e., the total benefit must be greater than total control implies that the environmental quality is between E and E1.

Another condition is that the  $MBC = MCC$ . This occurs where both the curves intersect each other at the point  $OE_1$ . Again at any other point whether  $MBC$  will be greater than Environmental Quality  $MCC$  or  $MCC$  will be greater than  $MBC$ . So only at  $OE_1$  the difference between  $TBC$  and  $TCC$  are maximum and the  $MCC = MBC$ . This occurs because at 'ab' the difference between  $TCC$  and  $TBC$  is maximum.

Figure 4: showing the optimum level of environmental quality

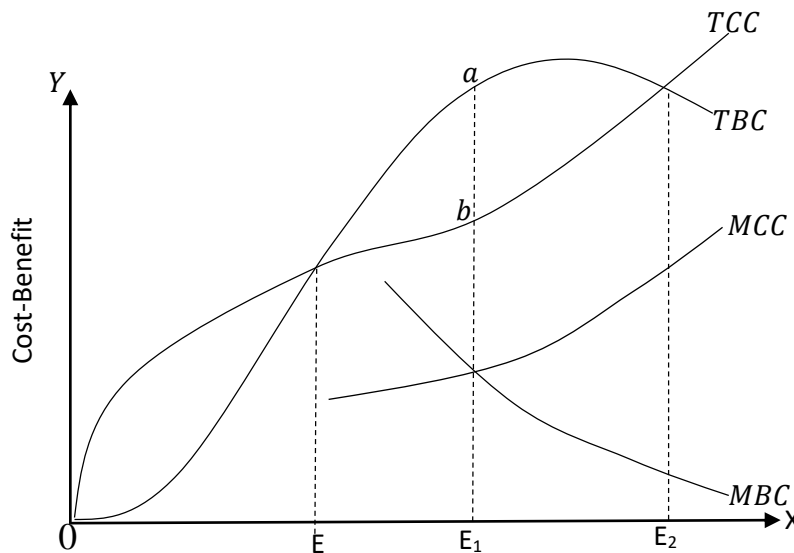


Fig. 3.26: Environmental Quality

Source: Wash, (2017). *Principles and Practice of Public Sector*

### 3.8. Cost Benefit Analysis—The Frame Work:

Knowing the cost benefit in theoretical form it is important to know it in terms of monetary terms. This can be studied under the property price approach. The property price approach is studied on the basis of noise pollution, this noise pollution is applicable to waste water, air pollution, etc.

**3.8.1. Property price approach:** The property price approach states that the people can buy peace and quiet by choosing their house at a quiet place. According to property price approach there are three types of movers, they are:

- (i) Natural movers.
- (ii) Movers due to noise.
- (iii) Bearer whatever the noise may be.

Natural movers are people those who move due to other factors, may be the job, education etc. They keep on moving their house from one place to the other. The next is the people who move due to the nuisance created by the noise. These type of people move in search of calm and quietness. Another type of people are those who bear the noise and never shift their place. This may be due to poverty or due to their love in ancestral property. However, keeping this in mind we can make some assumptions on this approach:

- (i) Individuals are free to choose the house according to their will.
- (ii) Noise is not common but is present in some areas.
- (iii) Peaceful area is available in plenty for the people to choose their dwelling.
- (iv) Noise or calmness is measurable and quantifiable as the commodities.

With these assumptions we can see the illustration with the help of a diagram.

In the diagram (Fig. 5.) X-axis denotes the number of housing units and the Y-axis the price paid for it. DH is the demand curve for houses and the stock of houses is assumed

to be fixed at ON. So the NH is the supply curve and the DH is the demand curve which intersects at H. the market price with the demand and supply curve is OP. Now consider that MN houses are affected by noise so the demand curve falls to the position  $D_1H_1$ . This leads to the fall in price of  $P_1$  for a noisy house.

Figure 5: showing housing unit to the Price paid by households

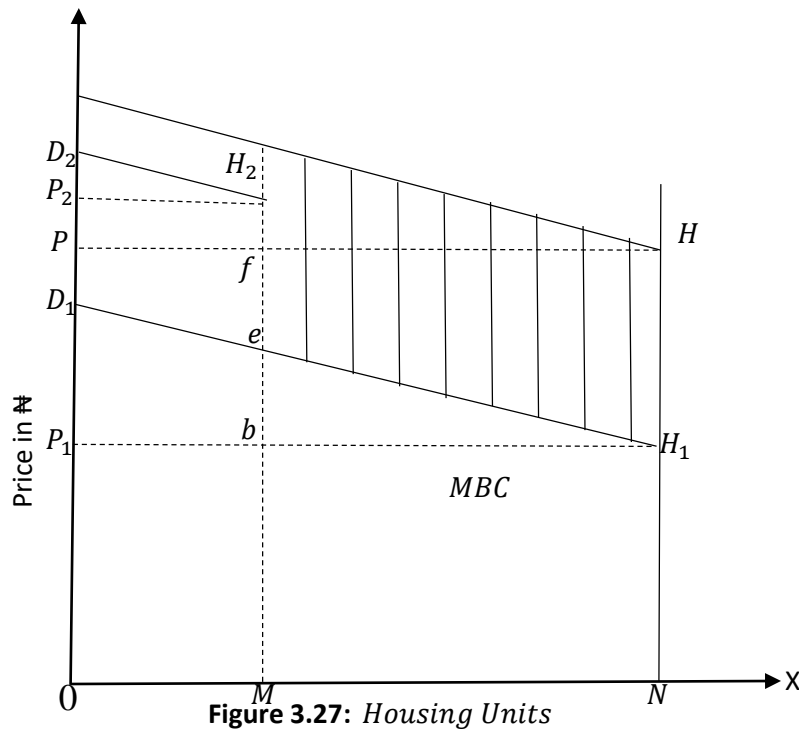


Figure 3.27: Housing Units

**Source: Wash, (2017). Principles and Practice of Public Sector**

Now OM quite houses are more valuable so the demand curve shifts to  $D_2H_2$ . This position is fixed higher because people are willing to pay more for the quite houses. This is given by the distance  $ae = D$ . the distance between the demand curves  $DH$  and  $D_1H_1$  for marginal consumer of quite houses.

If this marginal willing to pay is not altered, the marginal consumer must be willing to pay  $P_1 + D$  for quite houses. Hence adding  $D$  to  $P_1$  gives point  $H_2$  as the demand by the marginal consumer and similar analysis for other consumers gives the demand curve  $D_2H_2$  as the demand curve for quite houses setting a price of  $P_2$ .

According to the figure the observed house price differential after the introduction of noise concept is  $P_2 - P_1$  But the actual change in welfare is measured as the change in consumer surplus is given by the shaded area. This shaded area can be analysed as

$$aHH_1b = afH + feH_1H - ebH_1$$

Giving the symbol 'S' for consumers' surplus, for the surplus at the original price  $P$  and  $S_i$  for the surplus at the new price of the noisy house  $P_1$  this becomes.

$$\hat{S} = S_o + (P - P_1) MA - S_i \text{ (or) } (P - P_1) MN + (S_o - S_i)$$

This reveals that the measure of welfare loss to the house price is the differential between the initial 'no noise' situation and the new price of noisy houses plus the change in surplus between the initial 'no noise' situations. But the formula we derived earlier

stated the difference between the new price for quiet houses and the new price for noisy prices. From the figure we can observe the following inequality.

$$(P - P_i) MN < S < (P_2 - P_1) MN$$

This means, the approach using only the difference between the no-noise and noise situation will understate noise cost and an approach using the differential between the new price of noisy houses will overstate noise costs.

### **3.9. Merits and Demerits of Cost Benefit Analysis:**

#### **Merits:**

- (i) The cost benefit analysis may be applicable for both the new as well as old projects.
- (ii) The cost benefit analysis is based of accepted social principle that is on individual preference.
- (iii) This method encourages development for new techniques for the evaluation of social benefits.

#### **Demerits:**

- (i) The government is not completely aware of all the costs and benefits associated with the programme.
- (ii) This approach does not clearly state that who should bear the pollution control costs.
- (iii) The method of collecting data for this analysis is generally biased.
- (iv) The people will have different value system and there will always be loser in the process.

### **SELF-ASSESSMENT EXERCISE**

Discuss the principles of benefit and cost.

### **4.0. CONCLUSION**

A cost-benefit analysis is a process by which business decisions are analyzed. The benefits of a given situation or business-related action are summed, and then the costs associated with taking that action are subtracted. Some consultants or analysts also build the model to put a dollar value on intangible items, such as the benefits and costs associated with living in a certain town, and most analysts will also factor opportunity cost into such equations.

### **5.0. SUMMARY**

In this unit, we have learnt that reducing the positive and negative impacts of a project to their equivalent money value Cost-Benefit Analysis determines whether on balance the project is worthwhile. The equivalent money value is based upon information derived from consumer and producer market choices; that is, the demand and supply schedules for the goods and services affected by the project. Care must be taken to properly allow for such things as inflation. When all this has been considered a worthwhile project is one for which the discounted value of the benefits exceeds the discounted value of the costs; i.e., the net benefits are positive. This is equivalent to the

benefit/cost ratio being greater than one and the internal rate of return being greater than the cost of capital.

#### **6.0. TUTOR-MARKED ASSIGNMENT**

1. Define the term “Cost Benefit Analysis”
2. Discuss the steps in cost benefit analysis
3. Explain the differences between total benefit curves and marginal benefit curves analysis
4. List and explain the merits and demerits of cost benefit analysis
5. Discuss the analysis of the property price approach

#### **7.0. REFERENCES/FURTHER READING**

Adewole, A. F., (2016). *Introduction to Public Sector Economics*, 2nd edition, Ibadan Press Limited.

Yayah, E. T., (2014). *Introduction to Public Sector Economics*, 1st edition, Mac and Mix Publication limited.

Wash, W. Y. (2017). *Principles and Practice of Public Sector*, 1st edition, Macgraw, Publication limited.

## **MODULE 4: WELFARE ECONOMICS AND LINEAR PROGRAMMING**

### **UNIT 1: SOCIAL WELFARE**

#### **CONTENTS**

- 1.0 Introduction
- 2.0 Objectives
- 3.0 Main Content
  - 3.1 Criteria of Social Welfare
  - 3.2 Social Welfare Maximisation
  - 3.3 Individualistic Social Welfare Functions
- 4.0 Conclusion
- 5.0 Summary
- 6.0 Tutor-Marked Assignment
- 7.0 References/Further Reading

#### **1.0 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 task of welfare economics is to (1) show that in the present state,  $W < W^*$ , and (2) to seeks ways of raising  $W$  to  $W^*$ .

#### **2.0 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 maximisation

#### **3.0 MAIN CONTENT**

##### **3.1 Criteria of Social Welfare**

To evaluate alternative economic situation, we need some criteria 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, i.e, the growth of GNP, as a welfare criterion. According to him economic growth results in the increase of social welfare because growth increases employment, and the goods available for consumption to the society. To Adam Smith, economic growth means bringing  $W$  closer  $W^*$ .



This criterion however implies the acceptance of existing pattern of income distribution. Moreover, growth may lead to a reduction in social welfare, depending on who benefits mostly from it.

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. This implies that welfare is the sum of the utilities of the individuals of the society.

$$\text{Thus } W = U_A + U_B + U_C$$

According to Bentham,  $W > 0$  if  $(dU_A + dB_B + dU_C) > 0$ . In this arrangement it may be that  $U_A$  and  $U_B$  increase while  $U_C$  decreases. In other words, two individuals are better off while the third is worse off after the change in utility has taken place, but the sum of the increases in A and B is greater than the decrease in the utility of C.

Bentham's criterion implies that A and B have a greater worthiness than C. That is, it implies interpersonal comparison of the deservingness of the members of the society. Another difficulty is that the criterion cannot be applied to compare situations where "the greatest good and the members" do not exist simultaneously.

### **C. Cardinalist Criterion**

This advocates equality in the distribution of income and the argument is based on diminishing marginal utility of money. The more equal are income the greater is social welfare.

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 utility, 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 maintains that social welfare would be maximised if income was equally distributed to all members of the society.

This criterion has a number of shortcomings. It assumes that all individuals have identical utility functions for money, so that with equal distribution of income all would have the same marginal utility of money. This is too strong an assumption. 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. Moreover, equal distribution of income may adversely affect resource allocation. Redistribution may induce some people to work less, thus leading to an allocation of resources which produces a smaller total output.

## D. The Pareto-Optimality Criterion

According to this criterion any change which makes at least one individual better-off and no one worse-off is an improvement in social welfare.

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).

### I. *Efficiency of distribution of commodities among consumers.*

Let us assume that there are only two commodities, X and Y, produced with fixed amount of labour and capital, L and K, and two individuals, A and B, each with an ordinal preference function of the form.

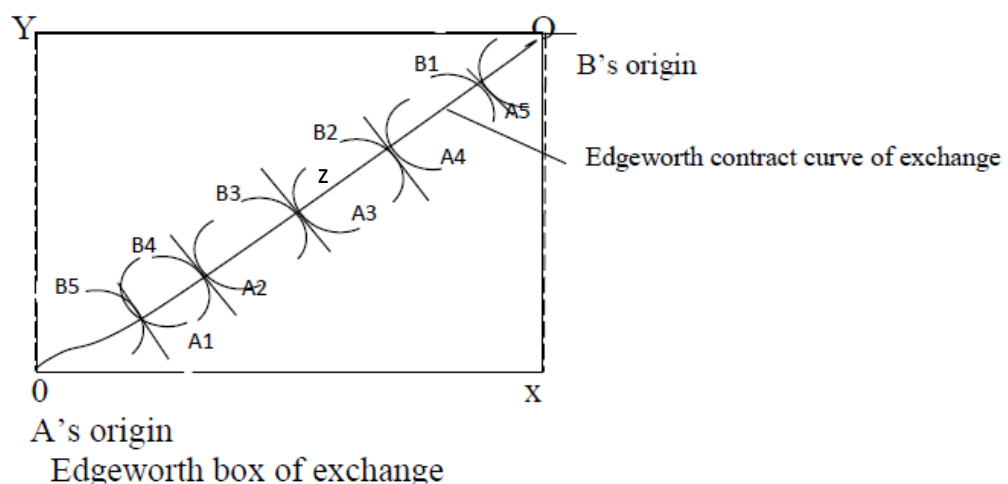
$$U_A = U_A (X_A Y_A)$$

$$U_B = U_B (X_B Y_B)$$

These are assumed to yield convex indifference curves of the usual type.

Given these assumptions we may then present the distribution of x and y between A and B in the Edgeworth box diagram as follow:

The joint equilibrium of production of the two firms in our simple model can be derived by use of the Edgeworth box of production.



**Figure 3.28: Joint Equilibrium of production for two firms**

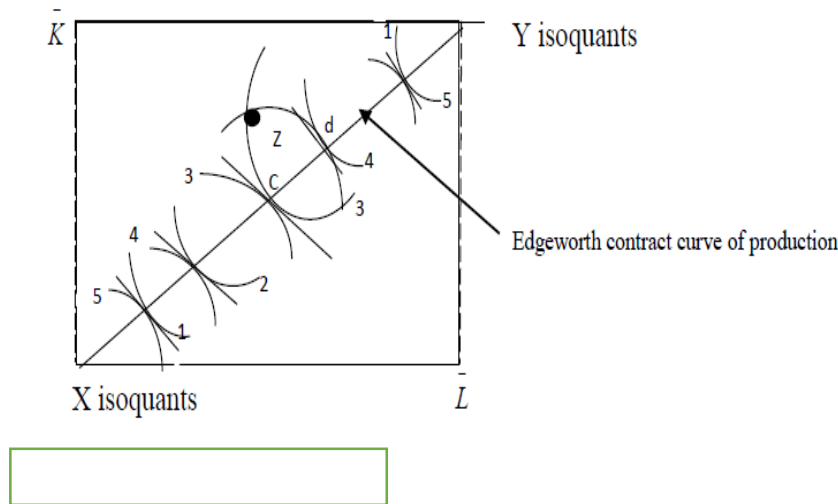
Distribution at  $z$  is off the contract curve and so is inefficient for example.  $B$  is indifferent between points  $z$  and  $b$ . Similarly,  $A$  is indifferent between  $z$  and  $a$ . moving from  $z$  towards  $a$  would improve the welfare of  $B$  without worsening that of  $A$  and similarly for moving from  $z$  towards  $b$ . Thus, for distribution to be efficient it must take place along the contract curve. So the contract curve shows the locus of Pareto optimal or efficient distribution of goods between consumers. This curve is a locus of points of tangency of the two sets of indifference curves, i.e., the points where the slopes of the indifferent curves are equal, satisfying the condition.

$$MRS_{X,Y}^A = MRS_{X,Y}^B \dots \dots \dots (1)$$

Therefore, the marginal condition for a Pareto-optimal distribution of commodities among consumers requires that the MRS between two goods to equal for all consumers.

**II. Efficiency of allocation of factors among producers:-**

To derive the Pareto-optimal condition for the allocation of factors among producers we also use the box diagram



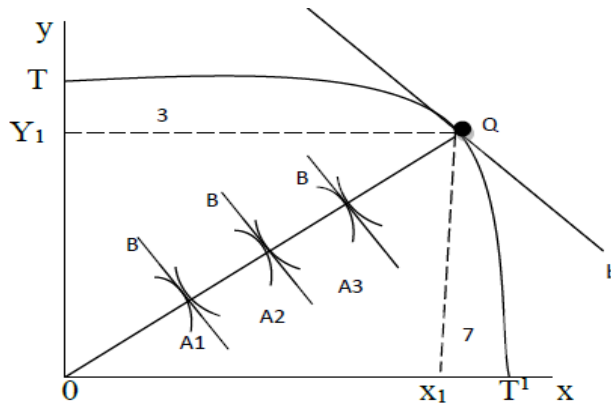
**Figure 3.29: Edgeworth Box of Production**

Given  $L$  and  $K$ , point  $Z$  in the box diagram is inefficient, since a reallocation of the factors between the producers of  $x$  and  $y$  such as to reach any point from  $c$  to  $d$  inclusive results in the increase of at least one commodity without a reduction in the other. The contract curve is a locus of points of tangency of the isoquants of the two firms which produce  $X$  and  $Y$ . Thus, at each point of the contract curve the following condition holds.

$$MRTS_{L,K}^X = MRTS_{L,K}^Y \dots \dots \dots (2)$$

This refers to the marginal condition for Pareto-optimal allocation of factors which requires that the MRTS between labour and capital be equal for all commodities produced by different firms.

**III. Efficiency in the composition of output (product mix)**



**Fig 3.30:** Edgeworth box of product mix

In Figure 3.30, TT' represents the production possibility curve or frontier, referred to as the production transformation curve between the two commodities X and Y. The slope of the transformation curve represents the marginal rate of transformation,  $dy/dx$ , of Y into X. For each point on the transformation schedule TT such as a or b, it is possible to construct an Edgeworth box with the dimension of X and Y combination at the point.

The efficient distribution of this output combination requires that the individuals involved be on the contract curves that represents the locus of efficient distribution.

The marginal condition for a Pareto-optimal or efficient composition of output requires that the marginal rate of product transformation MRPT between any two commodities be equal to the MRS between the same two goods :

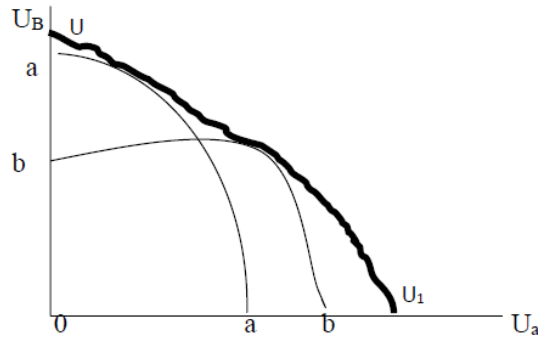
$$\text{Thus: } MRST_{x,y} = MRS_{X,y}^A = MRS_{X,y}^B \dots \dots \dots (2.3)$$

In summary, Pareto-optimal state in the economy can be attained if the following three marginal conditions are fulfilled.

- (1) The  $MRS_{x,y}$  between any two goods be equal for all consumers.
- (2) The  $MRTS_{L,K}$  between any two inputs be equal in the production of all commodities.
- (3) The MRPT be equal to the  $MRS_{x,y}$  for any two goods.

A situation may be Pareto-optimal without maximizing social welfare. However, welfare maximization is attained only in a situation that is Pareto-optimal. In other words, Pareto optimality is a necessary but not sufficient condition for welfare maximization.

To prove this, we first of all construct the grand utility possibility frontier. To achieve this we transcribe the points on the contract curve in terms of their utility indices on a pair of utility index axes, and this gives utility possibility curves for points such as a and b in fig. 3, because there are an infinite number of such points on TT', there will be an infinite number of utility possibilities curves such as aa and bb in fig. 4



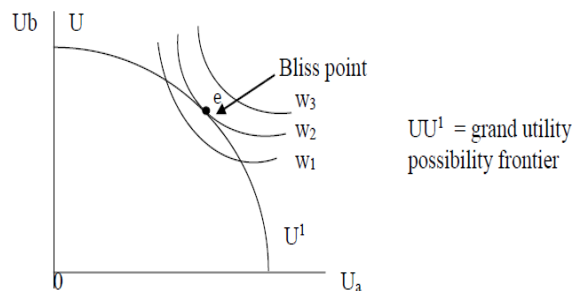
**Fig 3.31:** The Grand Utility Possibility Frontier.

We then find the envelope curve of this infinite number of utility possibility curves, such as  $UU_1$  in fig. 4. This is the grand utility possibility frontier. Each utility possibility  $UU_1$  curve will contribute one point on this frontier and this point will be found at that point on the contract curve where the equalized scope of the indifference curves, the equalized marginal rate of substitution, is equal to the scope of the transformation curve at the point defining the output combination for which the box is drawn.

Thus every point on the grand utility possibility frontier is Pareto-optimal. In other words, the frontier is a locus of optimal efficient points, but it does not give a unique solution to our quest for a point of welfare maximum.

Such a point can only be chosen through a social value judgement concerning the desirability of the welfare states of the individuals involved. We now have a resource to a social welfare function, from which we generate a set of social indifference curves,  $w_1, \dots, w_n$ , with all the properties of indifference curves, as shown in Fig. 5

The tangency of the grand utility possibility frontier with the highest possible social indifference curve gives the “bliss point” a uniquely defined welfare maximum for the society.



**Figure 3.32:** Maximization of social welfare.

### Weaknesses of Pareto Criterion

1. The criterion cannot evaluate a change that makes some people better-off and others worse-off. It therefore has limited applicability in the real world since most government policies involve changes that benefit some individuals at the expense of others. For instance, a measure that succeeds in protecting domestic industries would benefit entrepreneurs who reap higher profits and the employees but would make the lot of consumers who now pay higher prices worse.

A Pareto-optimal situation does not guarantee social welfare maximum. We have shown earlier that the grand utility possibility frontier is a locus of points that are

Pareto-optimal but the decision as to which of these points yield maximum social welfare requires the formation of the social welfare function and hence social indifference curves.

### **E. The Kaldor-Hicks Compensation Criterion**

Nicholas Kaldor and John Hicks proposed the 'compensation criterion' which states that a change in social welfare constitutes an improvement if those who benefit from it could compensate those who are hurt, and still be left with some 'net gain'. This criterion is illustrated in terms of money as follows: assume that a change in the economy will benefit some (gainers) and hurt others (losers). One can ask the gainers how much money they are prepared to pay for the change and ask the losers how much money they are prepared to pay in order to prevent the change from taking place. If the amount of money the gainers are prepared to pay is greater than that the losers are prepared to pay, then it implies that the gainers can compensate the losers for their loss and still have some 'net gain'. In this case, the change constitutes an improvement in social welfare.

This criterion, however, evaluates alternative situations on the basis of monetary valuations of different persons. Thus it implicitly implies that the marginal utility of money is the same for all individuals in the society. But this is not true, given that the distribution of income is not equal in the society. A given amount of money is marginally worth less to a rich man than to a relatively poor man. Thus, if the relatively richer segment of the society constitutes the gainers while the relatively poorer section constitutes the losers, a given amount of money, say N200.00 may be worth less to the gainers than, say N100.00 to the losers. So the fact that the gainers can compensate the losers and still have a net gain of N100.00 does not guarantee an improvement in social welfare because total utility would be reduced by the change since the disutility of the losers exceeds the utility of the gainers. Thus, only if the marginal utility of money is equal for all individuals would the Kaldor-Hicks criterion be a correct measure of welfare.

### **SELF-ASSESSMENT EXERCISE**

Identify the different criteria of social welfare

#### **Social Welfare Maximisation**

The condition for social welfare maximisation in the simple two factors, two commodities, and two consumer model is 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 maximisation and the goal of firms is profit maximisation.

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 maximisation?

### **3.2 Individualistic Social Welfare Functions**

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:

$$W = W(u_1(x_1), \dots, u_n(x_n)).$$

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 categorisation 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.

### **4.0 CONCLUSION**

From our discussion so far, we can infer that:

- Welfare functions are often used by economists of one sort or another to represent distributional judgments about allocations.
- 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 maximising some welfare function.

### **5.0 SUMMARY**

The concept of social welfare was reviewed, with extensive discussion on different criteria of social welfare. Pareto efficient allocation and the criteria emphasise the

allocative efficiency in the society. Different conditions for social welfare maximisation were also discussed before individualistic social function was reviewed.

## **6.0 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. Analyse the concept of social welfare function.

## **7.0 REFERENCES/FURTHER READING**

Koutsoyiannis, A. (2001). *Modern Microeconomics*, (2nd ed.). Hampshire: Macmillan Press Ltd.

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

Lipsey, et al. (2008). *Economics*, (13th ed.). Upper Saddle River, NJ: Pearson.

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



## **UNIT 2: EXTERNALITIES**

### **CONTENTS**

- 1.0 Introduction
- 2.0 Objectives
- 3.0 Main content
  - 3.1 Meaning of Definition of Externalities
  - 3.2 Types of Externality
  - 3.3 Effects of Externalities
  - 3.4 Externalities and Allocative Efficiency
  - 3.5 Solving the Externality Problem
- 4.0 Conclusion
- 5.0 Summary
- 6.0 Tutor-Marked Assignment
- 7.0 References/Further Readings

### **1.0. INTRODUCTION**

The principle that voluntary exchange benefits both buyers and sellers is a fundamental building block of the economic way of thinking. But what happens when a voluntary exchange affects a third party who is neither the buyer nor the seller? As an example, consider a club promoter who wants to build a night club right next to your apartment building. You and your neighbours will be able to hear the music in your apartments late into the night. In this case, the club's owners and attendees may both be quite satisfied with their voluntary exchange, but you have no voice in their market transaction. The effect of a market exchange on a third party who is outside or "external" to the exchange is called an externality. Because externalities that occur in market transactions affect other parties beyond those involved, they are sometimes called spillovers.

Almost all externalities are considered to be technical externalities. These types of externalities have an impact on the consumption and production opportunities of unrelated third parties, but the price of consumption does not include the externalities. This makes it so there is a difference between the gain or loss of private individuals and the aggregate gain or loss of the society as a whole. Oftentimes the action of an individual or organization results in positive private gains but detracts from the overall economy. Many economists consider technical externalities to be market deficiencies. This is why people advocate for government intervention to curb negative externalities through taxation and regulation.

Most externalities are negative. Pollution, for example, is a well-known negative externality. A corporation may decide to cut costs and increase profits by implementing new operations that are more harmful for the environment. The corporation realizes costs in the form of expanding its operations but also generate returns that are higher than the costs. However, the externality also increases the aggregate cost to the economy and society, making it a negative externality. Externalities are negative when the social costs outweigh the private costs.

Some externalities are positive. Positive externalities occur when there is a positive gain on both the private level and social level. Research and development (R&D) conducted by a company can be a positive externality. R&D increases the private profits of a company but also has the added benefit of increasing the general level of knowledge within a society. So, while a company such as Google profits off of its Maps application, society as a whole greatly benefits in the form of a useful GPS tool. Positive externalities have public, or social, returns that are higher than the private returns.

## 2.0. Objectives

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

- ✓ Define and understand the meaning of Externalities
- ✓ Know the types of Externalities
- ✓ Know the effects of Externalities and Allocative Efficiency
- ✓ Understand how to solve the problem of Externality.

## 3.0. Main Content

### 3.1. Meaning and Definition of Externalities

Externalities occur because economic agents have effects on third parties that are not parts of market transactions. Examples are: factories emitting smoke and dust, jet planes waking up people, or loudspeakers generating noise. These activities are all having a direct effect on the well-being of others that is outside direct market channels.

In these cases market prices (of soaps, air travel and entertainment) may not accurately reflect social cost because they take no account of the damage being done to third parties. Information being conveyed by the prices is fundamentally inaccurate, leading to a misallocation of resources.

An externality occurs whenever the activities of one economic agent affect the activities of another agent in ways that do not get reflected in market transactions. This is why externalities are taken as examples of market failure.

### 3.2. Types of Externalities

Externalities are of different types. Here we consider four main types of externality

#### (i) Inter Firm (Production) Externalities:

Suppose there are two firms in the economy — firm I is producing X and firm II is producing Y. Each uses only single input, labour. The production of good Y is said to have an external effect on the production of X if the output of X depends not only on the amount of labour chosen by firm I but also on the level at which the production of Y is carried on. In this case the production function for good X can be expressed as

$$X = f(L_x, Y) \dots \dots \dots (1)$$

where  $L_x$  denotes the amount of labour used to produce and Y indicates an effect on production over which firm I has no control. Negative inter-firm (or firm-firm) externality exists if  $\frac{\partial X}{\partial Y} < 0$ , i.e., increase in output of Y causes less of X to be produced.

#### (ii) Beneficial Externalities:

The activity of one firm may also have beneficial effect on others. For example, if a power plant is set up near a coal mine, hopefully, more coal can be extracted due to an

abundant supply of power. In this case,  $\frac{\partial X}{\partial Y} > 0$ . However, in the usual perfectly competitive case, the productive activities of one firm have no direct effect on those of other firms:  $\frac{\partial X}{\partial Y} = 0$ .

**(iii) Externalities in Utility (Consumption Externalities):**

Externalities also can occur if the activities of an economic agent directly affect an individual's utility. Most obvious examples are environmental externalities (such as noise from a loud radio). Such externalities may sometimes be beneficial.

(Mr. John may actually enjoy the song being played on Mr. Sen's radio) this type of externality arises when one individual's utility depends directly on the utility of someone else. If, for example, Mr. A cares about Mr. B's welfare, we can express A's utility

$$U_A = f(X_1, X_2, \dots, X_n; U_B) \dots \dots (2)$$

where  $X_1, X_2, \dots, X_n$  are the goods which A consumes and  $U_B$  is B's utility.

If A wants B to be better-off (if A were a close relative of B,  $\frac{\partial U_A}{\partial U_B}$  would be positive. If, on the other hand, A were envious of B,  $\frac{\partial U_A}{\partial U_B}$  would be negative; that is improvements in B's utility make A worse off. If A were indifferent to B's welfare,  $\frac{\partial U_A}{\partial U_B} = 0$ .

**(iv) Public Goods Externalities:**

Public goods or collective consumption goods (such as national defence, roads, bridges, public parks, public school, hospitals, etc.) create externality problems because such goods can be allocated through the market and those who enjoy such goods do not pay prices directly.

The cost of providing such goods is covered through taxes. Once such goods are produced (either by the government or by some private agency) they provide benefits to all the members of society. This is because such goods are to be consumed jointly. It is not possible to restrict these benefits to the specific group of individuals who pay for them.

So the benefits are available to all. For example, once a national defence system is established, all individuals in society are protected by it whether they wish to be or not and whether they pay for it or not. Choosing the right level of output for such a good is a complex task, since market signals are not quite accurate.

**3.3. Effects of Externalities**

Externalities create divergence between social benefit and private benefit and between social cost and private cost. In the presence of positive externality, marginal social benefit (of any activity such as education or health/medical care) = marginal private benefit + marginal external benefit. This is why, in the presence of positive externality, a commodity or service is under produced its actual output is less than the socially desirable level. And there is need to subsidies all activities which generate positive externalities and cause departure from Pareto optimality.

Likewise, in the presence of negative externalities, marginal social cost = marginal private cost + marginal external cost. As result a commodity or service is overproduced.

Actual output exceeds the socially desirable level, the activity generating negative externality has to be taxed in order to ensure Pareto optimality.

### 3.4. Externalities and Allocative Efficiency

The presence of externalities can cause a market to operate inefficiently. This point may now be illustrated. Let us assume that two firms are located near each other and that one of these (II) has negative effect on the production of the other (I). Suppose the production function of the firm II which generates pollution is expressed as

$$Y = g(L_Y) \dots \dots \dots (3)$$

where  $L_Y$  = the quantity of labour devoted to the production of Y. The production function for good X (which exhibits an externality) was given by equation (1). The Pareto conditions for an optimal allocation of labour require that the social marginal revenue product of labour (SMRP<sub>L</sub>) be equal for both firms. If  $P_x$  and  $P_y$  are the prices of good X and good Y, respectively, the SMRP<sub>L</sub> in the production of good X is given by

$$SMRP_L^x = P_x \frac{\partial f}{\partial L_x} \dots \dots \dots (4)$$

Due to the presence of production externality, the statement of the SMRP<sub>L</sub><sup>x</sup> in the production of Y is more complex. An extra unit of labour employed by firm II will produce some extra Y.

But it will also generate some extra pollution, and will reduce the output of AT, produced by Firm I, Consequently,

$$SMRP_L^y = P_y \frac{\partial g}{\partial L_y} + P_x \frac{\partial f}{\partial Y} \cdot \frac{\partial Y}{\partial L_y} \dots \dots \dots (5)$$

where the second term on the right hand side represents the effect of hiring additional workers in the production of K on the value of production of X. This effect will be negative if  $\frac{\partial f}{\partial Y} < 0$  Efficiency then requires that

$$SMRP_L^y = SMRP_L^x \dots \dots \dots (6)$$

Independent decision-making by the two firms will normally not ensure the fulfilment of this condition. Firm 1 (producing X) will hire labour up to the point at which its private MRP, is equal to the prevailing wage rate

$$w = MRP_L^x = P_x \frac{\partial f}{\partial L_x} \dots \dots \dots (7)$$

Firm II with follow suit, and we have

$$w = MRP_L^y = P_y \frac{\partial g}{\partial L_y} \dots \dots \dots (8)$$

The market will, therefore, equate private marginal revenue products, but this market equilibrium will ensure Pareto efficiency only if  $\frac{\partial f}{\partial Y} = 0$  in equation (5). In other words in the presence of externalities, the decisions of the two firms or their managers will not bring about an optimal allocation.

Since we have assumed that  $\frac{\partial f}{\partial Y} < 0$ , labour will be over allocated to the production of good Y. The  $SMRP_L$  in the production of Y will fall short of that in the production of X. If, on the other hand, we assume that  $\frac{\partial f}{\partial Y} > 0$ , then labour will be under-allocated to the production of Y.

### 3.5. Solving the Externality Problem

There are certain solutions to the allocation problems posed by externality. Two such solutions are taxation and merger:

**1. Taxation:** The government can impose a suitable excise duty on the firm generating the external diseconomy. This tax is likely to cause the output of Y to be cut back and would cause labour to be shifted out of the production of Y. This standard remedy was first suggested A. C. Pigou in the 1920s and is known as the Pigouvian tax.

The taxation solution is illustrated in Fig. 1. The demand for Y is given by  $D_Y$  and the private marginal cost curve for Y by  $MC$ . The curve  $MC'$  shows the social marginal cost of production of Y. Thus the socially optimal level of output is  $Y_2$ . However, in the presence of negative externalities, the normal functioning of the market will cause output level of  $Y_1$  to be produced.

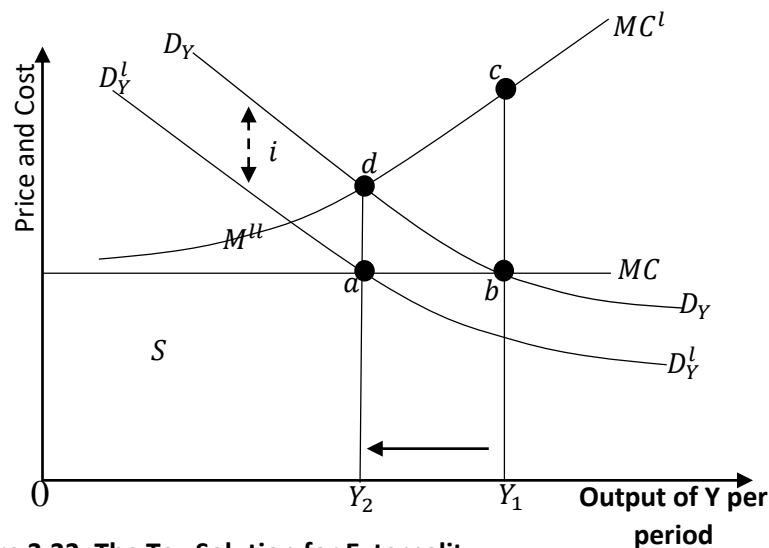


Figure 3.33: The Tax Solution for Externality

One way to force the market to allocate goods correctly would be to impose an excise duty of  $t$  per unit of Y produced. The effect of this indirect tax is to shift the demand curve facing the firm from  $D_Y$  to  $D_Y^l$  and this will cause the profit maximizing the level of output of Y to fall from  $Y_1$ , to  $Y_2$ . This is a government solution to the externality problem.

**2. Merger and Internalization:** A private solution for the allocation of distortions caused by the externality between X and Y would be for the two firms to merge. If a single firm operates both plants X and Y, it will recognise the harmful effect that production of Y has on the production function for good X.

In effect, the new (merged) firm would now bear the full social marginal costs of Y production because it also produces X now. In other words, the firm would now take

the marginal cost curve for Y production to be  $MC^1$  in Fig. 1 and would produce at the point where

$$P_y = MC^1 \dots\dots\dots (9)$$

which is exactly what is required for allocative efficiency.

The externality in the production of Y has been internalized as a result of the merger. The reason is that what was marginal external cost before the merger has now become a part of the marginal private cost of the merged firm.

**3. Defining property rights:** The stricter definition of property rights can limit the influence of economic activities on unrelated parties. However, it is not always a viable option since the ownership of particular things such as air or

**4. Subsidies:** A government can also provide subsidies to stimulate certain activities. The subsidies are commonly used to increase the consumption of goods with positive externalities.

### **SELF-ASSESSMENT EXERCISE**

Discuss the term “Externalities”.

### **4.0. CONCLUSION**

In this unit, we conclude that an externality is a consequence of an economic activity experienced by unrelated third parties; it can be either positive or negative. Pollution emitted by a factory that spoils the surrounding environment and affects the health of nearby residents is an example of a negative externality. The effect of a well-educated labour force on the productivity of a company is an example of a positive externality.

Economic production can cause environmental damage. This trade-off arises for all countries, whether they be high-income or low-income, and whether their economies are market-oriented or command-oriented.

An externality occurs when an exchange between a buyer and seller has an impact on a third party who is not part of the exchange.

An externality can have a negative or positive impact on the third party. If those parties imposing a negative externality on others had to take the broader social cost of their behaviour into account, they would have an incentive to reduce the production of whatever is causing the negative externality.

In the case of a positive externality, the third party is obtaining benefits from the exchange between a buyer and a seller, but they are not paying for these benefits. If this is the case, markets tend to under-produce output because suppliers do not consider the additional benefits to others. If the parties that are creating benefits for others can somehow be compensated for these external benefits, they would have an incentive to increase production.

### **5.0. SUMMARY**

In this unit, we have learnt and discuss on meaning and definition of externalities, types of externality, effects of externalities, externalities and the Allocative efficiency and how to solve the externality problem.

## **6.0. TUTOR-MARKED ASSIGNMENT**

1. Discuss the causes of Externalities
2. List the private solutions that can be used to deal with externalities. Explain in detail including proper graphs, diagrams and figures.
3. Discuss the effects of externalities and how do we solve Externalities problems?

## **7.0. REFERENCES/FURTHER READING**

Adesanjo, W. R., (2014). Impact of Externalities on the health of the Communities in Nigeria, *a Journal Ibadan Printing Press limited*, Vol 4, pp. 22-34.

Volt, J. W., (2016). Externalities and Manufacturing Industry, *A Journal Macgrill Publication Limited*, Vol 9, pp. 45-57.

Wright, G. H., (2015). *Principles and Practice of Finance* (1<sup>st</sup> ed.), Pat and Park Publication limited.

## **UNIT 3: ECONOMIC EFFICIENCY AND EQUITY; SOCIAL AND PRIVATE COSTS.**

### **CONTENTS**

- 1.0 Introduction
- 2.0 Objectives
- 3.0 Main Content
  - 3.1 Efficiency
  - 3.2 Efficiency and Equity
  - 3.3 Private Versus Social Costs
  - 3.4 Externalities
- 4.0 Conclusion
- 5.0 Summary
- 6.0 Tutor-Marked Assignment
- 7.0 References/Further Reading

### **1.0 INTRODUCTION**

Most people recognise that resources are scarce and that factors of production must be used wisely. If resources are wasted, fewer goods and services will be produced and fewer wants and needs can be satisfied. Economic decision making must be efficient so that benefits gained are greater than costs incurred. On the other hand, many people, for example, believe in equal pay for equal work. As a result, it is illegal to discriminate on the basis of age, sex, race, religion, or disability in employment. When it comes to selling products, most people feel that advertisers should not be allowed to make false claims about their products.

### **2.0 OBJECTIVES**

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

- Understand the concept of economic efficiency
- Understand the efficiency and equality mix
- Differentiate between private and social cost
- Understand what externalities entails

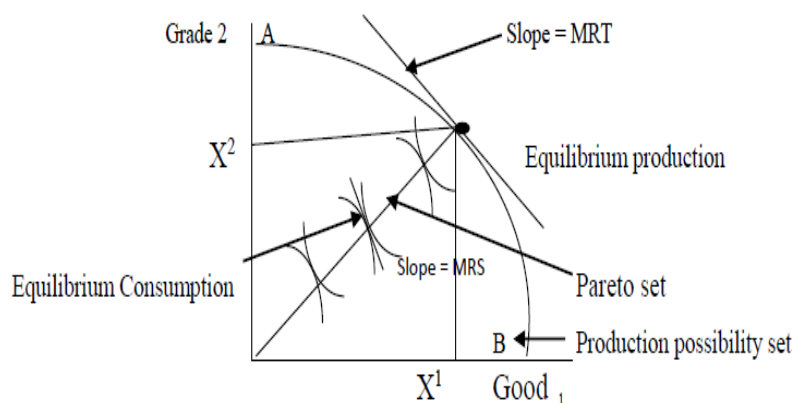
### **MAIN CONTENT**

#### **3.1 Efficiency**

Ordinarily efficiency means getting any given results with the smallest possible inputs or getting the maximum possible output from given resources. For example, efficiency in consumption means allocating goods between consumers in such a way that it would not be possible by any reallocation to make some people better-off without making anybody else worse off. Again efficiency in production means allocating the available resources between industries so that it would not be possible to produce more of some goods without producing less of any others. Efficiency in the choice of the set of goods to produce means choosing this set so that it would not be possible to change it so as to make some consumers better off without others becoming worse off. Efficiency is also referred to as Pareto Optimality.



The relationship between efficiency and Pareto Optimality can be demonstrated as follows:



**Fig 3.34:** Production and the Edgeworth box.

In Figure 3.34, line AB represents the production possibility frontier in a given economy denoting the combination of goods 1 and 2 that could be produced if all resources are fully employed  $X_1$  and  $X_2$  indicate aggregate consumption levels of goods 1 and 2. The Slope = MRS Pareto efficient consumption levels will lie along the Pareto set (the Edgeworth Contract line) – the line of mutual tangencies of the indifference curves as illustrated in the figure (1). These are the allocations in which each consumer’s marginal rate of substitution – the rate at which he or she is just willing to trade – equals that of the other.

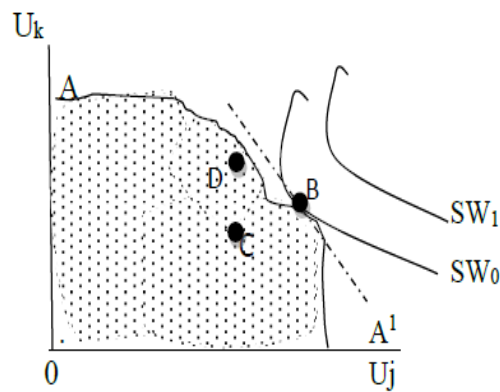
These allocations are Pareto efficient as far as the consumption decisions are concerned. If people can simply trade one good for another, the Pareto set describes the set of boundless that exhausts the gains from trade. But in an economy with production, there is another way to “exchange” one good for another – namely to produce less of one good and more of another - that is at another point on the production possibility frontier.

The Pareto set in the above example describes the set of Pareto efficient boundless given the amounts of goods 1 and 2 available. In other words, in an economy with production those amounts can themselves be chosen out of the production possibilities set.

### 3.2 Efficiency and Equity

Decisions made on the grounds of efficiency may clash with those made on considerations of equity or fairness. For example, efficiency suggests that it is not rational to tax other consumer goods but exempt food, while equity suggests exempting food from tax as the poor spend relatively more of their incomes on food.

Let us use Figure 2 for an illustration.



**Fig 3.35: Social Allocation of utility**

In the figure, AA represents the degree of social satisfaction or utility frontier. The dotted area therefore represents the social opportunity set. If the society is made up of only two individuals, Karl and John, their respective utilities are shown on the vertical and horizontal axes

The overall social welfare maximum is at point B where the highest attainable social welfare function is tangent to the total utility frontier. Suppose that a social choice is to be made between alternative allocations D and C. D is in the Pareto-efficient set and C is not. But the distribution between the two parties concerned (Karl and John) is such that D is not Pareto preferred to C, Karl is better off at D but John is worse off.

Decision on the basis of efficiency would dictate allocation at B. What then are the factors responsible for the choice of D or C? The answers to this question may be:

- (b) The desirable social outcomes such as the goals of policies such as the reduction in income inequality.
- (c) The satisfaction of desirable wants of individuals.
- (d) Societal values which do not accord with efficiency criterion, such as what is fair, just, equitable, etc, in the eyes of the society.

### 3.3 Private Versus Social Costs

Private cost is the cost of providing goods or services as it appears to the persons or firm supplying them. This includes the cost of any factor services or inputs bought by the supplier, the value of the work done, and the use of land, buildings, and equipment owned by the supplier. Private cost excludes any external harm caused to other people or the environment, such as noise or pollution, unless the supplier is legally obliged to pay for it.

Social cost, on the other hand, is the total cost of any activity. This includes not only private costs which fall directly on the person or firm conducting the activity, but also external costs which fall on other people who are not able to exact any compensation for them.

### 3.4 Externalities

Let us illustrate the relationship between private and social costs from the point of view of production externalities. Suppose that firms produce some amount of steels, and also produces a certain amount of pollution,  $x$ , which it dumps into a river. Firm F, a fishery,

is located downstream and is adversely affected by S's pollution. Suppose that firm S's cost function is given by

$$C_s(S, x), \dots \dots \dots (3.4.1)$$

Where s is the amount of steel produced and x is the amount of pollution produced Firm F's cost function is given by  $C_f(f, x)$ ; f indicates the production of fish and x the amount of pollution. (Note that F's costs of producing a given amount of fish depend on the amount of pollution produced by the steel firm). We will suppose that pollution increases the cost of providing fish that is  $\frac{dc_f}{dx} > 0$  and that pollution decreases the cost of steel production,  $\frac{dc_s}{dx} > 0$ . This last assumption says that increasing the amount of pollution will decrease the cost of steel production and that reducing pollution will increase the cost of steel production, at least over some range. The steel firm profit-maximization problem is to maximise:

$$Px_s - C_s(s, x) \dots \dots \dots (3.4.2)$$

and the fishery's profit-maximization problem is to maximise:

$$P_f F - c_f(f, x) \dots \dots \dots (3.4.3)$$

Note that the steel mill gets to choose the amount of pollution that it generates, but the fishery firm must take the level of pollution as outside its control.

The conditions characterizing profit maximization will be

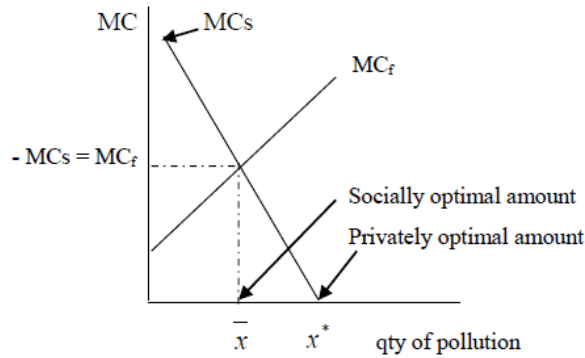
$$P_s = \frac{dc_s(s^*, x^*)}{ds} \dots \dots \dots (3.4.4)$$

$$0 = \frac{dc_s(s^*, x^*)}{ds} \dots \dots \dots (3.4.5)$$

for the steel firm and  $P_f = \frac{dc_f(f^*, x^*)}{df} \dots \dots \dots (3.4.6)$

for the fishery. Those conditions say that at the profit-maximizing point, the price of each good – steel and pollution – should equal its marginal cost. In the case of steel firm, one of its products is pollution, which by assumption, has a zero price. So the condition determining the profit maximizing supply of pollution says to produce pollution unit the cost of an extra unit is zero.

It is not hard to see the externality here: the fishery cares about the production of pollution but has no control over it. The steel firm looks only at the cost of producing steel when it makes its profit-maximizing calculations. It does not consider the cost it imposes on the fishery. The increase in the cost of fishing associated with an increase in pollution is part of the social cost of steel production, and it is being ignored by the steel firm. In general, we expect, that the steel firm will produce too much pollution from a social point of view since it ignores the impact of that pollution on the fishery.



**Fig 3.36:** Social cost and private cost.

The steel firm produces pollution up to the point where the marginal cost of extra pollution equals zero. But the Pareto efficient production of pollution is at the point where price equals marginal social cost, which includes the cost of pollution borne by the fishery.

#### 4.0 CONCLUSION

From our discussion so far, we can infer that Economic efficiency is a more general concept that occurs when any change that benefits someone would result in harm for someone else. Note that technical efficiency is a necessary condition for economic efficiency since a movement toward the production possibilities curve would benefit one or more individuals and we also illustrated the relationship between private and social costs from the point of view of production externalities.

#### 5.0 SUMMARY

We saw that efficiency is also referred to as Pareto optimality and private cost excludes any external harm caused to other people or the environment, such as noise or pollution, unless the supplier is legally obliged to pay for it while Social cost, on the other hand, is the total cost of any activity.

#### 6.0 TUTOR-MARKED ASSIGNMENT

What determines economic efficiency as described by Vilfredo Pareto.

#### 7.0 REFERENCES/FURTHER READING

- Koutsoyiannis, A. (2001). *Modern Microeconomics* (2nd ed.). Hampshire: Macmillan Press Ltd.
- Krugman, P. & Robin, W. (2012). *Economics*. (3rd ed.). New York: Worth Publishers.
- Lipsey, et al. (2008). *Economics*. (13th ed.). Upper Saddle River, NJ: Pearson.
- Robert, S. P. & Daniel, L. R. (2009). *Microeconomics*. (7th ed.). New Jersey: Pearson Education International.

## **UNIT 4: LINEAR PROGRAMMING CONTENTS**

- 1.0 Introduction
- 2.0 Objectives
- 3.0 Main Content
  - 3.1 Concept of Linear Programming
  - 3.2 Assumptions of Linear Programming
  - 3.3 Advantages and Limitations
  - 3.4 Construction of the Model
  - 3.5 Methods of Solving Linear Programming Problems
  - 3.6 The Simplex Method In Practice
- 4.0 Conclusion
- 5.0 Summary
- 6.0 Tutor-Marked Assignment
- 7.0 References/Further Reading

### **1.0 INTRODUCTION**

Linear programming deals with the optimization (maximization or minimization) of a function of variables known as the objective functions. It is subject to a set of linear equalities and/or inequalities known as constraints. Linear programming is a mathematical technique which involves the allocation of limited resources in an optimal manner, on the basis of a given criterion of optimality.

In this unit, properties of Linear Programming Problems (LPP) are discussed. The graphical method of solving LPP is applicable where two variables are involved. The most widely used method for solving LPP problems consisting of any number of variables is called simplex method, developed by G. Dantzig in 1947 and made generally available in 1951.

### **2.0 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
- ✓ Solve maximization and minimization problems using graphical, algebra or simplex methods.

### **3.0 MAIN CONTENT**

#### **3.1 Concept of Linear Programming**

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:

$$\begin{aligned} & \text{Minimize } c^T x \\ & \text{Subject to } Ax \leq b \\ & \text{Maximize } x \geq 0 \end{aligned}$$

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  $(.)^T$  is the matrix transpose. The expression to be maximised or minimised is called the objective function ( $c^T x$  in this case). The inequalities  $Ax \leq b$  and  $x \geq 0$  are the constraints which specify a convex polytope over which the objective function is to be optimized. 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 utilized for some engineering problems. Industries that use linear programming models include transportation, energy, telecommunications, and manufacturing. It has proved useful in modelling diverse types of problems in planning, routing, scheduling, assignment, and design.

Linear programming maximization problem looks like the equation presented below:

$$\begin{aligned} \text{Max } f(x) &= cx \\ \text{Subject to } Ax &\leq b, x \geq 0. \end{aligned}$$

$$\begin{aligned} \text{More generally,} \\ \text{Max } Z &= P_1 X_1 + P_2 X_2 \end{aligned}$$

$$\begin{aligned} \text{Subject to} \\ a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n &\leq b_1 \\ a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n &\leq b_2 \\ a_{m1}x_1 + a_{m2}x_2 + \dots + a_{mn}x_n &\leq b_m \\ x_1 \geq 0, x_2 \geq 0, \dots, x_n &\geq 0. \end{aligned}$$

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).

In linear programming  $z$ , the expression being optimized is called the objective function. The variables  $x_1, x_2 \dots x_n$  are called decision variables, and their values are subject to  $m + 1$  constraints (every line ending with a  $b_i$ , plus the non-negativity constraint). A set of  $x_1, x_2 \dots x_n$  satisfying all the constraints is called a feasible point and the set of all such points is called the feasible region. The solution of the linear program must be a point  $(x_1, x_2 \dots x_n)$  in the feasible region, or else not all the constraints would be satisfied.

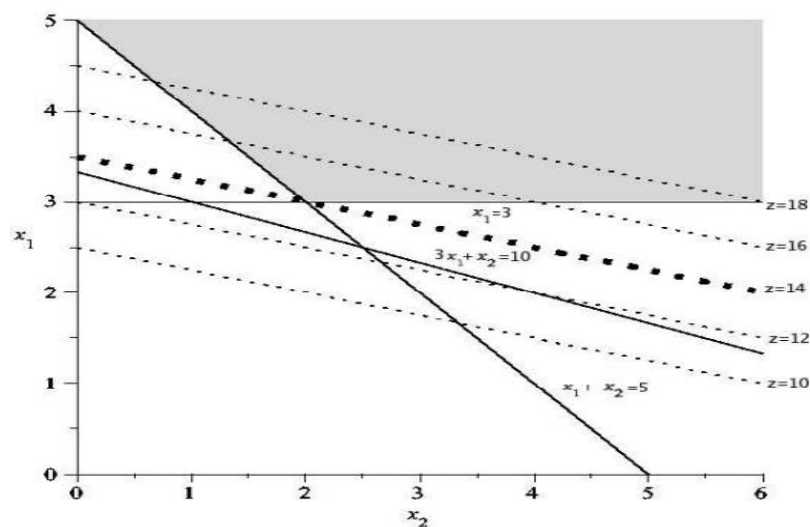
The following example illustrates that geometrically interpreting the feasible region is a useful tool for solving linear programming problems with two decision variables. The linear program is:

$$\begin{aligned} \text{Minimize } 4x_1 + x_2 &= z \\ \text{Subject to} \\ 3x_1 + x_2 &\geq 10 \\ x_1 + x_2 &\geq 5 \\ x_1 &\geq 3 \\ x_1, x_2 &\geq 0. \end{aligned}$$

We plotted the system of inequalities as the shaded region in Figure below. Since all of the constraints are “greater than or equal to” constraints, the shaded region above all three lines is the feasible region. The solution to this linear program must lie within the shaded region.

Recall that the solution is a point  $(x_1, x_2)$  such that the value of  $z$  is the smallest it can be, while still lying in the feasible region. Since  $z = 4x_1 + x_2$ , plotting the line  $x_1 = (z - x_2)/4$  for various values of  $z$  results in isocost lines, which have the same slope.

Along these lines, the value of  $z$  is constant. In the Figure below, the dotted lines represent isocost lines for different values of  $z$ . Since isocost lines are parallel to each other, the thick dotted isocost line for which  $z = 14$  is clearly the line that intersects the feasible region at the smallest possible value for  $z$ . Therefore,  $z = 14$  is the smallest possible value of  $z$  given the constraints. This value occurs at the intersection of the lines  $x_1 = 3$  and  $x_1 + x_2 = 5$ , where  $x_1 = 3$  and  $x_2 = 2$ .



In the shaded region in the figure above, all three solid lines is the feasible region (one of the constraints does not contribute to defining the feasible region). The dotted lines are isocost lines. The thick isocost line that passes through the intersection of the two defining constraints represents the minimum possible value of  $z = 14$  while still passing through the feasible region.

### 3.2 Assumptions of LP for Production

Before we get too focused on solving linear programs, it is important to review some theory. For instance, several assumptions are implicit in linear programming problems.

These assumptions are:

- ✓ Certainty: This assumption is also called the deterministic assumption. This means that all parameters (all coefficients in the objective function and the constraints) are known with certainty. Realistically, however, coefficients and parameters are often the result of guess-work and approximation.
- ✓ The objective function is linear
- ✓ There is an additivity of resources and activities. The contribution of any variable to the objective function or constraints is independent of the values of the other variables.
- ✓ Activities and resources are non-negative

- ✓ Resources and activities are infinitely divisible. Decision variables can be fractions. However, by using a special technique called integer programming, we can bypass this condition. Unfortunately, integer programming is beyond the scope of this paper.
- ✓ There is a linear relationship between activities and resources
- ✓ Linear programming models are deterministic and not stochastic.
- ✓ Proportionality The contribution of any variable to the objective function or constraints is proportional to that variable. This implies no discounts or economies to scale. For example, the value of 8x1 is twice the value of 4x1, no more or less.

### 3.3 Advantages and Limitations

#### Advantages

- It helps decision-makers to use their productive resources effectively.
- The decision-making approach of the user becomes more objective and less subjective.
- 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. A significant advantage of linear programming is highlighting such bottle necks.

#### Limitations

- 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.
- Factors such as uncertainty and time are not taken into consideration.
- Parameters in the model are assumed to be constant but in real life situations, they are not constant.
- Linear programming deals with only single objectives, whereas in real life situations, we may have multiple and conflicting objectives.
- 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?

### 3.4 Construction of the Model

You assume that a business firm produces two commodities, X and Y, with two different inputs, Labour (L) and Capital (K). The total quantities of L and K available per unit of time are specified as  $L = 1600$  labour hours; and  $K = 2000$  units. In addition, assume that producing 1 unit of commodity X requires 4 units of labour (L) and 2 units of capital (K). One unit of commodity Y requires 2 units of L and 5 units of K. Profits per unit of commodities X and Y are estimated at N10 and N8, respectively.

Given these information, the firm's objective is to maximise its total profit ( $\pi$ ). The problem is to choose an output mix of X and Y that maximises profit.

The following steps are needed

#### ***1. Transformation of the Problem into Linear Programming***

To transform the problem at hand to linear programming format, you need to restate the conditions of the problem in programming language.

Take note of the following steps, they will help you.



**Step 1:** Specification of the Objective Function. The firm's objective function can be expressed in the following form:

$$\text{Maximise } \pi = 10X + 8Y \dots \dots \dots (3.4.1)$$

where X and Y represent quantities of commodities X and Y. When you multiply these quantities by their unit prices (or profits as the case may be) you will obtain the total profit ( $\pi$ ) as indicated by equation (3.4.1).

It will be the linear programming technique that you will use in determining the units of X and Y to produce in order that profit will be maximised.

**Step 2:** Specification of the relevant Constraint Inequalities. Using the information available in the question, you will formulate the relevant constrain equations as:

The constraint inequality for input L may be specified thus,

$$4X + 2Y \leq 1600 \dots \dots \dots (3.4.2)$$

The constraint inequality for input K may similarly be specified as,

$$2X + 5Y \leq 2000 \dots \dots \dots (3.4.3)$$

**Step 3:** Specification of Non-negative Conditions. Note that a negative quantity in optimum solutions is not allowed and does not make economic sense so that, you must impose non-negative conditions in the linear programming problem. The relevant non-negative conditions for the problem at hand can be expressed as:

$$X \geq 0 \text{ and } Y \geq 0 \dots \dots \dots (3.4.4)$$

You are now in a position to formulate the required linear programming problem in terms of equations and inequalities. The problem becomes:

$$\text{Maximise } \pi = 10X + 8Y \text{ (the objective function),}$$

*Subject to the constraints:*

$$4X + 2Y \leq 1600 \dots \dots \dots (3.4.5)$$

$$2X + 5Y \leq 2000, \dots \dots \dots (3.4.6)$$

*where  $X \geq 0$  and  $Y \geq 0$*

You will obtain the optimum solution to the problem at hand by solving for the values X and Y in the above equations.

### 3.5 Methods of Solving Linear Programming Problems

You will now be introduced to two popular methods of solving linear programming problems:

- (i) Graphical Method
- (ii) Simplex Method

#### 3.5.1 The Graphical Method

This method is the simplest technique in solving linear programming problem. You begin by converting the constraint inequalities into equalities, and then sketching them in a graph. Thus, the constraint inequality (3.4.5) becomes:

$$4X + 2Y = 1600 \dots \dots \dots (3.5.1)$$

*and that of (3.4.6) becomes:*

$$2X + 5Y = 2000 \dots \dots \dots (3.5.2)$$

Notice that equations (3.5.1) and (3.5.2) are linear equations in X and Y.

To sketch these equations, you will begin by determining the intercept terms for the two-dimensional graph in X and Y. Thus, to graph equation (3.4.1), you obtain the Y- and X-axis as,

**For the Y – axis:**

$$4X + 2Y = 1600$$

$$2Y = 1600 - 4X$$

$$\text{When } X = 0, 2Y = 1600$$

$$Y = 800$$

The Y – intercept is therefore, 800.

**For the X – axis:**

$$4X + 2Y = 1600$$

$$4X = 1600 - 2Y$$

$$\text{When } Y = 0, 4X = 1600$$

$$X = 400$$

The X – intercept is therefore, 400.

Similarly, for equation (3.5.2), you obtain the Y- and X-axis as:

$$2X + 5Y = 2000$$

$$5Y = 2000 - 2X$$

$$\text{When } X = 0, 5Y = 2000$$

$$Y = 400$$

Here, the Y – intercept is 400, and,

$$2X + 5Y = 2000$$

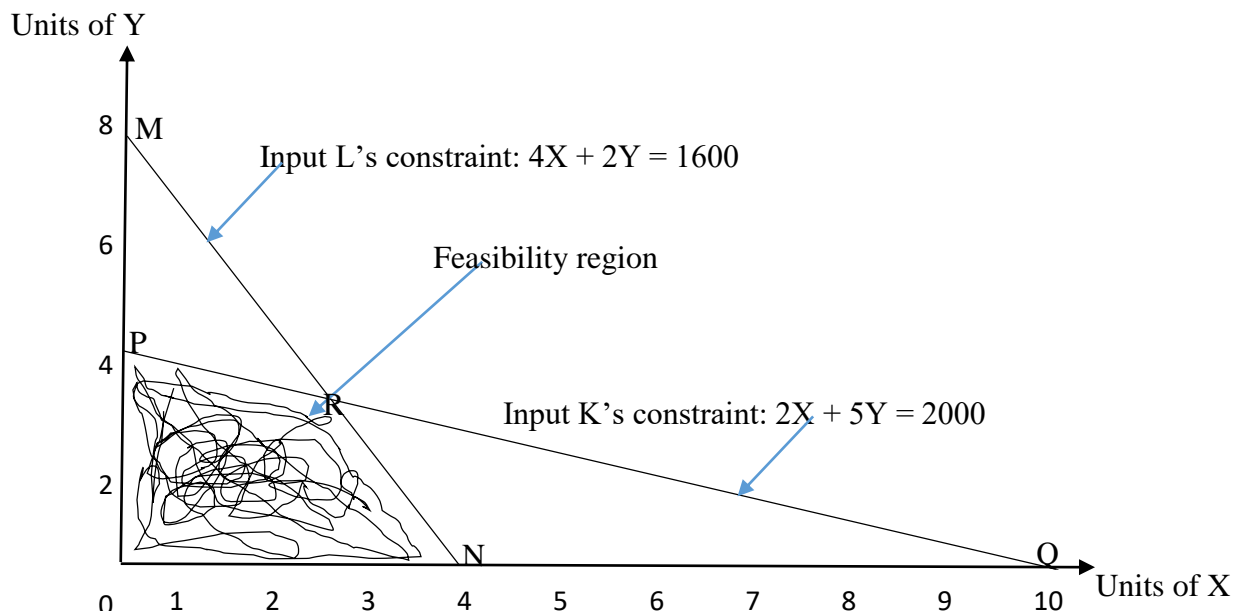
$$2X = 2000 - 5Y$$

$$\text{When } Y = 0, 2X = 2000$$

$$X = 1000$$

Therefore, the X – intercept is 1000.

The sketches are as in Figure 3.5.1 below:



**Fig. 3.37: Production Constraints and Feasibility Region**

Observe that in Figure 3.37, the line MN is formed by joining the Y and X-intercepts for the labour (L) constraint equation, and that of PQ is formed by joining Y and X-intercepts for the capital (K) constraint equation. All the points on the line MN satisfy the constraint,  $4X + 2Y \leq 1600$ . The area under OMN is referred to as the feasibility

space for the single input, L. This implies that any point within the feasibility space and on the border lines is a feasible point for this input.

Similarly, the area under OPQ is referred to as the feasibility space for the single input, K. All the points on the line PQ satisfy the constraint,  $2X + 5Y \leq 2000$ .

The shaded area under OPRN represents the feasible region, where you will obtain the feasible output choices. Each of these choices satisfies both the constraints and the stated non-negativity conditions. Only those points falling under the feasible region satisfy all the feasibility conditions. Any point to the right of the area marked MRP represents a combination of the commodities X and Y that cannot be produced within the limited availability of the inputs L and K. All the points marked by the area PMR satisfy only the constraint,  $4X + 2Y \leq 1600$ .

Similarly, all the points within the area marked NRQ satisfy only the constraint,  $2X + 5Y \leq 2000$ . It follows that only the feasible area, OPRN that meets the constraints and contains the point of solutions to the profit maximisation problem.

Your next step is to locate the point on the boundary of the feasible area.

This point will represent the combination of the commodities X and Y that maximises profit. You can do this by graphing or sketching the objective function in the form of isoprofit lines, for different output levels, and superimposing these over the feasible region.

Graphing of the objective function would require finding the slope of the objective function, which you can do as follows:

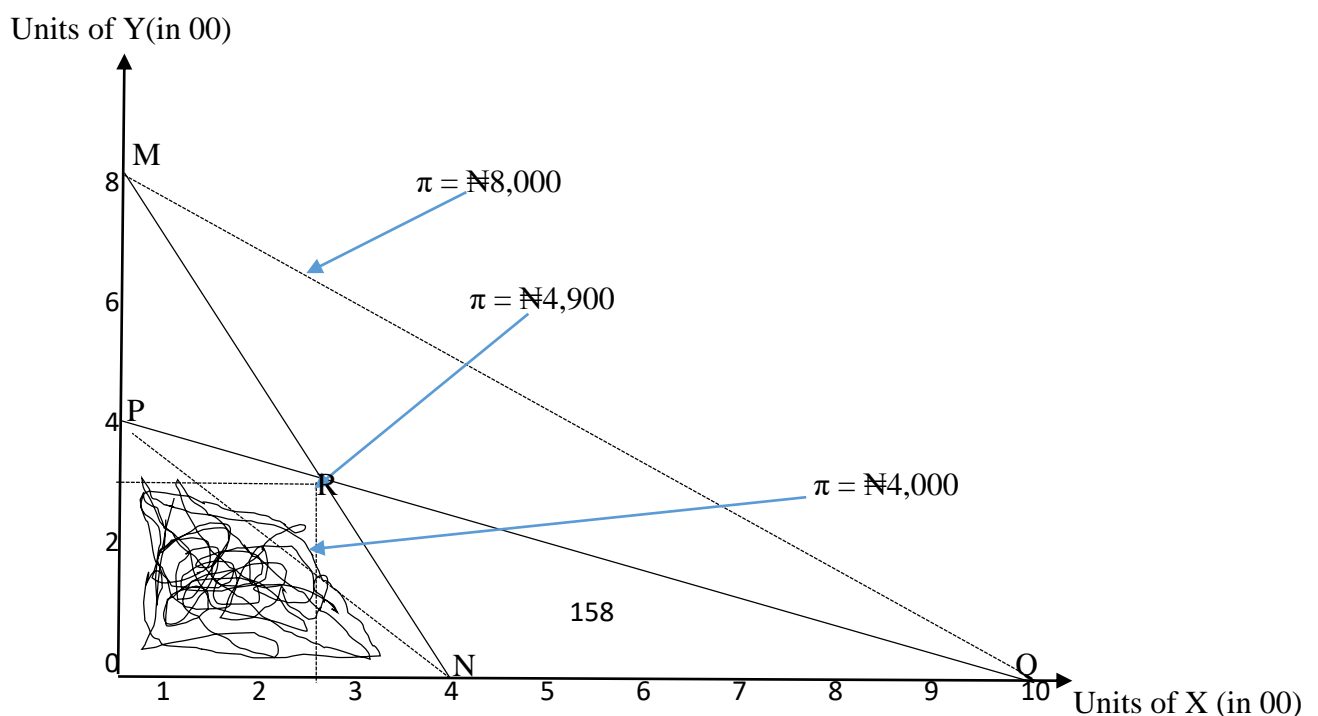
Given the objective function,  $\pi = 10X + 8Y$ , you write it in terms of Y to get:

$$Y = \frac{\pi}{8} - \frac{10}{8}X = \frac{\pi}{8} - 1.25X$$

where  $\pi = 0$ ,

$Y = -1.25X$ .

The coefficient, -1.25 gives the slope of the isoprofit line. This means that 1.25 units of commodity Y would yield the same profit as 1 unit of X. With this slope, a series of isoprofit lines can be drawn and superimposed over the feasibility region, as you can observe in Figure 3.5.2 below. Note that since profitability of the two commodities is constant, isoprofit lines are parallel to each other.



Observe that the isoprofit line marked N8,000 is not possible because it lies very much above the feasibility space or region. The isoprofit line marked N4,000, passing under the feasibility space reflects an underutilisation of inputs. It therefore, indicates a less than maximum profit. The shaded area to the right of this line indicates the scope for increasing profit. The isoprofit line marked N4,900, which is tangent to the boundary of feasibility space at point R is the highest possible isoprofit line, representing the maximum possible profit given the resource constraints. The tangential point, R which represents a combination of X and Y (that is,  $2.5(100) = 250$  units, and  $3(100) = 300$  units), yields the maximum profit. Thus, Maximum Profit ( $\pi^*$ ) =  $10(250) + 8(300) = \text{N}4,900$ .

It follows that the profit maximising units of commodities X and Y are 250 units and 300 units, respectively.

Alternatively

$$4X + 2Y = 1600 \dots\dots\dots (3.5.1)$$

$$2X + 5Y = 2000 \dots\dots\dots (3.5.2)$$

$$Y = 800 - 2X \text{ (making } Y \text{ the SF in equation 3.5.1)..... (3.5.3)}$$

$$2X + 5(800 - 2X) = 2000 \text{ (substituting eqn. 3.5.3 into eqn. 3.5.2)}$$

$$2X + 4000 - 10X = 2000$$

$$2X - 10X = 2000 - 4000$$

$$-8X = -2000$$

$$X = 250 \text{ units}$$

$$Y = 800 - 2(250) = 300 \text{ units}$$

It follows that the profit maximising units of commodities X and Y are 250 units and 300 units, respectively.

### SELF-ASSESSMENT EXERCISE

$$\min z = -3x_1 + 8x_2$$

$$\text{s. t. } 4x_1 + 2x_2 \leq 12$$

$$2x_1 + 3x_2 \leq 6$$

$$x_1, x_2 \geq 0$$

- (a) Sketch the feasible region in the  $(x_1, x_2)$  space.
- (b) Identify the regions in the  $(x_1, x_2)$  space where the slack variables  $s_1$  and  $s_2$ , you would have introduced, are equal to zero.
- (c) Solve the problem using graphical method.

### 3.6 The Simplex Method in Practice

The Simplex algorithm remedies the shortcomings of the aforementioned “brute force” approach. Instead of checking all of the extreme points in the region, the Simplex algorithm selects an extreme point at which to start. Then, each iteration of the algorithm takes the system to the adjacent extreme point with the best objective function value. These iterations are repeated until there are no more adjacent extreme points with better objective function values. That is when the system is at optimality.

The best way to implement the Simplex algorithm by hand is through tableau form. A linear program can be put in tableau format by creating a matrix with a column for each variable, starting with z, the objective function value, in the far left column. For a general linear program, of the form

$$\begin{aligned} & \text{Maximize } cx \\ & \text{Subject to } Ax \leq b \\ & \quad x \geq 0 \end{aligned}$$

the initial tableau would look like:

$z$	$-c$	$0$
$0$	$A$	$b$

The idea here is that  $z - cx = 0$  and  $Ax = b$  so in this format all the variables can be forgotten and represented only by columns. If we were actually completing the Simplex algorithm for this program, we would be concerned with the lack of an initial basis in the first tableau. If there is no initial basis in the tableau (an  $m \times m$  identity matrix), then we are not at an extreme point and the program cannot possibly be at optimality.

**A. Algebraic Simplex Method:** To begin with, here is a simple example.

**Illustration:**

Consider the product mix problem.

$$\begin{aligned} & \text{Maximize } z = 6x_1 + 5x_2 \\ & \text{Subject to } x_1 + x_2 \leq 5 \\ & \quad 3x_1 + 2x_2 \leq 12 \\ & \text{with } x_1, x_2 \geq 0 \dots \dots \dots (3.6.1) \end{aligned}$$

As in the last unit, you should first convert the inequalities to equations as shown below

$$\begin{aligned} & \text{Maximize } z = 6x_1 + 5x_2 + 0s_1 + 0s_2 \\ & \text{Subject to: } \quad x_1 + x_2 + s_1 + 0s_2 = 5 \\ & \quad 3x_1 + 2x_2 + 0s_1 + s_2 = 12 \\ & \text{with } x_1, x_2, s_1, s_2 \geq 0 \end{aligned} \tag{3.6.2}$$

Note that slack variables have zero contributions to the objective function, therefore, solving (3.6.1) is the same as solving (3.6.2).

One important thing about the Simplex method is that since it should not evaluate any infeasible solution, you would need to begin to solve with a basic feasible solution, and to do this, you will fix  $x_1 = 0$  and  $x_2 = 0$ , so that  $s_1 = 5$  and  $s_2 = 12$ .

In fact, one of the important things in any linear programming problem is that the constraints should not have a negative value on the right hand side. If the constraint has a negative value on the right hand side, then you will need to multiply the constraint by -1 to make it non-negative, although the sign of the inequality may be reversed. So you would make an assumption that all linear programming problem that you solve, the constraint should have a non-negative value on the right-hand side. It can have a zero but it should not have a negative.

Since each of these constraints have a non-negative value on the right hand side, and each of the slack variables appears in only one of the equations, it is now very easy to fix the rest of the variables to zero and have a starting solution of  $s_1 = 5$  and  $s_2 = 12$  which is basic feasible. It is basic because the variables  $x_1$  and  $x_2$  are fixed to zero, and

feasible because  $s_1$  and  $s_2$ , each appear only in one of the constraints and are non-negative.

Thus the first basic feasible solution (or the starting solution) for this problem is  $x_1 = x_2 = 0$ ,  $s_1 = 5$  and  $s_2 = 12$ .

### Iteration 1

Having identified the basic variables, i.e.  $s_1$  and  $s_2$ , write these basic variables and the objective function in terms of the non-basic variables  $x_1$  and  $x_2$ , as follows,  $s_1 = 5 - x_1 - x_2$ ,  $s_2 = 12 - 3x_1 - 2x_2$  and  $z = 6x_1 + 5x_2$ . If you set  $x_1 = 0$  and  $x_2 = 0$ , then  $s_1 = 5$ ,  $s_2 = 12$  and  $z = 0$ .

But you are interested in maximizing the objective function  $z$  which right now is zero, with  $x_1 = x_2 = 0$  and are non-basic. To increase  $z$ , you have to increase  $x_1$  and  $x_2$ , since both have strictly positive coefficients. In Simplex method, the idea is that you should increase one variable at a time. In other words, you will either increase  $x_1$  or  $x_2$  to maximize  $z$ . But since, the coefficient of  $x_1$ , 6 is greater than the coefficient of  $x_2$ , it is better to increase  $x_1$  because the rate of increase would be higher.

Presently,  $x_1 = 0$ . There will be a limit the value  $x_1$  can take because as you increase  $x_1$ , you will realize that  $s_1$  and  $s_2$  will decrease. For instance, if  $x_1 = 1$ ,  $x_2 = 0$  still, then  $s_1 = 4$  and  $s_2 = 9$ . Thus, as  $x_1$  increases,  $s_1$  and  $s_2$  start reducing to zero. Therefore you will increase to a point where one of them becomes zero, otherwise increasing  $x_1$  beyond that will end up making either  $s_1$  or  $s_2$  negative, which would violate the non-negativity restriction, and you do not want it.

Now looking at the equations and

$$s_1 = 5 - x_1 - x_2 \dots \dots \dots (3.6.3)$$

$$s_2 = 12 - 3x_1 - 2x_2 \dots \dots \dots (3.6.4)$$

The highest value  $x_1$  can take in (3.6.3) for  $x_2$  to remain non-negative is 5 and the highest it can take in (3.6.4) for  $x_2$  to remain non-negative is 4. So the highest value  $x_1$  can take is  $\min(5, 4) = 4$ . A further increase in  $x_1$  would result to a negative value of  $x_2$  and would violate the non-negativity restriction. Hence equation (3.6.4) becomes the binding equation which determines the highest value  $x_1$  can take. This leads us to the second iteration.

### Iteration 2.

Rewriting equation (3.6.4) for  $x_1$ , you will have and substituting in the rest gives you  $x_1 = 4 - \frac{2}{3}x_2 - \frac{1}{3}s_2$ ,  $s_1 = 5 - \left(4 - \frac{2}{3}x_2 - \frac{1}{3}s_2\right) - x_2 = 1 - \frac{1}{3}x_2 + \frac{1}{3}s_2$  and  $z = 6\left(4 - \frac{2}{3}x_2 - \frac{1}{3}s_2\right) + 5x_2 = 24 + x_2 - 2s_2$ .

In this iteration,  $x_1$  and  $s_1$  are basic, while  $x_2$  and  $s_2$  are non-basic. Letting  $x_2 = s_2 = 0$ , then  $x_1 = 4$ ,  $s_1 = 1$  and  $z = 24$ . This is another basic feasible solution that you have obtained. It is basic because  $x_2 = s_2 = 0$  and feasible because the value of the variables are non-negative.

Remember your objective is to increase  $z = 24 + x_2 - 2s_2$  further. This you can do by either increasing  $x_2$  or decreasing  $s_2$  ( $s_2$  has a negative coefficient). But  $s_2$  is non-basic and already at zero, so you cannot decrease  $s_2$  further, otherwise it will violate the non-

negativity restriction. Also  $x_2$  is non-basic and is zero, So you will increase  $x_2$  in order to increase  $z$ .

Consider the equations

$$x_1 = 4 - \frac{2}{3}x_2 - \frac{1}{3}s_2 \dots \dots \dots (3.6.5)$$

and

$$s_1 = 1 - \frac{1}{3}x_2 + \frac{1}{3}s_2 \dots \dots \dots (3.6.6)$$

you will observe that the highest value  $x_2$  can take in (3.6.5) so that  $x_1$  remains feasible is  $x_2 = 6$  and the highest value  $x_2$  can take in (3.6.6) so that  $s_1$  remains feasible is  $x_2 = 3$ . Thus, for both variables  $x_1$  and  $s_1$  to remain feasible, the highest value  $x_2$  can take is  $\min(6, 3) = 3$ . A further increase for the value of  $x_2$  beyond 3 makes  $s_1$  negative and would violate the non-negativity restriction. Hence equation (3.6.6) becomes the binding equation which determines the highest value  $x_2$  can take. This leads us to the third iteration.

### **Iteration 3**

Rewriting equation (3.6.6) for  $x_1$ , you will have and substituting in the rest gives you  $x_2 = 3 - 3s_1 + s_2$ ,  $x_1 = 4 - \frac{2}{3}(3 - 3s_1 + s_2) - \frac{1}{3}s_2 = 2 + 3s_1 - s_2$  and  $z = 24 + (3 - 3s_1 + s_2) - 2s_2 = 27 - 3s_1 - s_2$ .

In this iteration,  $x_1$  and  $x_2$  are basic, while  $s_1$  and  $s_2$  are non-basic. Letting  $s_1 = s_2 = 0$ , then  $x_1 = 2$ ,  $x_2 = 3$  and  $z = 27$ .

Now, you can check whether you can increase  $z = 27 - 3s_1 - s_2$  further. To increase  $z$  further, you can either decrease  $s_1$  or  $s_2$  because both have negative coefficients. But it is not possible to decrease any of  $s_1$  or  $s_2$  because both are already zero and decreasing them will make them infeasible. So you cannot proceed any further from this point to try and increase  $z$  further. Hence you will stop here and conclude that the best solution which is  $x_1 = 2$ ,  $x_2 = 3$  and  $z = 27$  have been obtained.

You will notice that this is the same solution that you can obtain with the graphical method when tried.

A close examination of this method shows you that you have done exactly the three important things you want it to do, which are

- It did not evaluate any infeasible solution because you put extra effort to determine the limiting value the entering variables can take so that the non-negativity restriction is not violated.
- It evaluated progressively better basic feasible solutions, because at each time you were only trying to increase the objective function for the maximization problem.
- It terminated immediately the optimum solution is reached. This is the simplex method represented in algebraic form.

### **B. Simplex Method-Tabular Form**

Here you will see the Simplex method represented in tabular form. The simplex method is carried out by performing elementary row operations on a matrix you would call the simplex tableau. This tableau consists of the augmented matrix corresponding to the

constraints equations together with the coefficients of the objective function written in the form

$$-c_1x_1 - c_2x_2 - \dots - c_nx_n + (0)s_1 + (0)s_1 + \dots + (0)s_m + z = 0$$

In the tableau, it is customary to omit the coefficient of z. For instance, the simplex tableau for the linear programming problem

$$\begin{aligned} \text{Maximize } z &= 4x_1 + 6x_2 \\ \text{Subject to: } &-x_1 + x_2 \leq 11 \\ &x_1 + x_2 \leq 27 \\ &2x_1 + 5x_2 \leq 90 \end{aligned} \dots\dots\dots(3.6.7)$$

By the addition of slack variables  $s_1, s_2$  and  $s_3$  to the constraints, you can rewrite the above problem as

$$\begin{aligned} \text{Maximize } z &= 4x_1 + 6x_2 + 0s_1 + 0s_2 + 0s_3 \\ \text{Subject to: } &-x_1 + x_2 + s_1 + 0s_2 + 0s_3 = 11 \\ &x_1 + x_2 + 0s_2 + s_2 + 0s_3 = 27 \\ &2x_1 + 5x_2 + 0s_1 + 0s_2 + s_3 = 90 \dots\dots\dots(3.6.8) \end{aligned}$$

Since slack variables have zero contributions to the objective function, solving (3.6.7) is the same as solving (3.6.8)

### Initial Simplex tableau

The initial simplex tableau for this problem is as follows

Table 3.1:

$B$	$x_1$	$x_2$	$S_1$	$S_2$	$S_3$	$x_B$
$S_1$	-1	1	1	0	0	11
$S_2$	1	1	0	1	0	27
$S_3$	2	5	0	0	1	90
$z_j - c_j$	-4	-6	0	0	0	0

For this initial simplex tableau, the basic variables are  $s_1, s_2$  and  $s_3$ , and the non basic variables (which have a value of zero) are  $x_1$  and  $x_2$ . Hence, from the two columns that are farthest to the right, you see that the current solution is

$$x_1 = 0, x_2 = 0, s_1 = 11, s_2 = 27, \text{ and } s_3 = 90$$

This solution is a basic feasible solution and is often written as

$$(x_1, x_2, s_1, s_2, s_3) = (0, 0, 11, 27, 90)$$

The entry in the lower-right corner of the simplex tableau is the current value of z. Note that the bottom-row entries under  $x_1$  and  $x_2$  are the negatives of the coefficients of  $x_1$  and  $x_2$  in the objective function

$$z = 4x_1 + 6x_2.$$

To perform an optimal check for a solution represented by the simplex tableau, you will look at the entries in the bottom row ( $z_j - c_j$  row) of the tableau. If any of these entries are negative (as above), then the current solution is not optimal.



## Pivoting

Once you have set up the initial simplex tableau for a linear programming problem, the simplex method consists of checking for optimality and then, if the current solution is not optimal, improving the current solution. (An improved solution is one that has a larger  $z$ -value than the current solution.) To improve the current solution, you will bring a new basic variable into the solution—you would call this variable the entry variable. This implies that one of the current basic variables must leave, otherwise you would have too many variables for a basic solution—you would call this variable the departing variable. You are to choose the entering and the departing variables as follows.

1. The entering variable corresponds to the smallest (the most negative) entry in the bottom (i.e.  $z_j - c_j$ ) row of the tableau.
2. The departing variable corresponds to the smallest non-negative ratio of  $b_i/a_{ij}$  in the column determined by the entering variable.
3. The entry in the simplex tableau in the entering variable's column and departing variable's row is called the pivot.

Finally, to form the improved solution, you will apply Gauss-Jordan elimination to the column that contains the pivot, as illustrated in the following example. (This process is called pivoting.)

### Pivoting to Find an Improved Solution.

Use the simplex method to find an improved solution for the linear programming problem represented by the following tableau.

**Table 3.2:**

$B$	$x_1$	$x_2$	$S_1$	$S_2$	$S_3$	$x_B$
$S_1$	-1	1	1	0	0	11
$S_2$	1	1	0	1	0	27
$S_3$	2	5	0	0	1	90
$z_j - c_j$	-4	-6	0	0	0	0

*The objective function for this program is  $z = 4x_1 + 6x_2$ .*

**Solution.** Note that the current solution ( $x_1 = 0, x_2 = 0, s_1 = 11, s_2 = 27, s_3 = 90$ ) corresponds to a  $z$ -value of 0. To improve this solution, you determine that  $x_2$  is the entering variable, because  $-6$  is the smallest entry in the  $z_j - c_j$  row.

To see why you should choose  $x_2$  as the entering variable, remember that  $z = 4x_1 + 6x_2$ . Hence, it appears that a unit change in  $x_2$  produces a change of 6 in  $z$ , whereas a unit change in  $x_1$  produces a change of only 4 in  $z$ .

To find the departing variable, you will locate the  $b_i$ 's that have corresponding positive elements in the entering variables column and form the following ratios

$$\theta := \frac{11}{1} = 11, \frac{27}{1} = 27, \frac{90}{5} = 18 \dots \dots \dots (3.6.9)$$

Here the smallest positive ratio is 11, so you will choose  $s_1$  as the departing variable.

**Table 3.3:**

$B$	$x_1$	$x_2$	$S_1$	$S_2$	$S_3$	$x_B$	$\theta$
$S_1$	-1	<b>1</b>	1	0	0	11	11
$S_2$	1	1	0	1	0	27	27
$S_3$	2	5	0	0	1	90	18
$z_j - c_j$	-4	6	0	0	0	0	

Note that the pivot is the entry in the first row and second column. Now, you will use Gauss-Jordan elimination to obtain the following improved solution.

$$\begin{array}{c} \text{Before} \\ \left[ \begin{array}{cccccc} -1 & 1 & 1 & 0 & 0 & 11 \\ 1 & 1 & 0 & 1 & 0 & 27 \\ 2 & 5 & 0 & 0 & 1 & 90 \\ -4 & -6 & 0 & 0 & 0 & 0 \end{array} \right] \end{array} \text{ vs } \begin{array}{c} \text{After} \\ \left[ \begin{array}{cccccc} -1 & 1 & 1 & 0 & 0 & 11 \\ 2 & 0 & -11 & 0 & 16 & 16 \\ 7 & 0 & -50 & 1 & 35 & 35 \\ -10 & 0 & 6 & 0 & 0 & 66 \end{array} \right] \end{array}$$

The new tableau now appears as follows

**Table 3.4:**

$B$	$x_1$	$x_2$	$S_1$	$S_2$	$S_3$	$x_B$
$S_1$	-1	1	1	0	0	11
$S_2$	2	0	-1	1	0	16
$S_3$	7	0	-5	0	1	35
$z_j - c_j$	-10	0	6	0	0	66

Note that  $x_2$  has replaced  $s_1$  in the basis column and the improved solution has a z-value of  $(x_1, x_2, S_1, S_2, S_3) = (0, 11, 0, 16, 35)$

$$z = 4x_1 + 6x_2 = 4(0) + 6(11) = 66$$

### Iteration 2

In the illustration above the improved solution is not yet optimal since the bottom row still has a negative entry. Thus, you can apply another iteration of the simplex method to further improve our solution as follows. You choose  $x_1$  as the entering variable. Moreover, the smallest non-negative ratio  $\frac{11}{-1} = -11$ ,  $\frac{16}{2} = 8$ , and  $\frac{35}{7} = 5$  is 5, so  $s_3$  is the departing variable. Gauss Jordan elimination produces the following.

**Table 3.5:**

$B$	$x_1$	$x_2$	$S_1$	$S_2$	$S_3$	$x_B$	$\theta$
$S_1$	-1	1	1	0	0	11	-
$S_2$	2	0	-1	1	0	16	8
$S_3$	<b>7</b>	0	-5	0	1	35	5
$z_j - c_j$	-10	0	6	0	0	66	

The pivot is the entry in the third row and the first column. Pivoting using Gaussian Elimination, you will obtain the following improved solution.

$$\left[ \begin{array}{cccccc} -1 & 1 & 1 & 0 & 0 & 11 \\ 2 & 0 & -11 & 0 & 16 & 16 \\ 7 & 0 & -50 & 1 & 35 & 35 \\ -10 & 0 & 6 & 0 & 0 & 66 \end{array} \right] \text{ vs } \left[ \begin{array}{cccccc} -1 & 1 & 1 & 0 & 0 & 11 \\ 2 & 0 & -11 & 0 & 16 & 16 \\ 1 & 0 & -\frac{5}{7} & 0 & \frac{1}{7} & 5 \\ -10 & 0 & 6 & 0 & 0 & 66 \end{array} \right]$$

$$\text{vs } \begin{bmatrix} 0 & 1 & \frac{2}{7} & 0 & \frac{1}{7} & 16 \\ 0 & 0 & \frac{3}{7} & 1 & -\frac{2}{7} & 6 \\ 1 & 0 & -\frac{5}{7} & 0 & \frac{1}{7} & 5 \\ 0 & 0 & \frac{8}{7} & 0 & \frac{10}{7} & 116 \end{bmatrix}$$

Thus, the new simplex tableau is as follows

**Table 3.6:**

$B$	$x_1$	$x_2$	$S_1$	$S_2$	$S_3$	$x_B$
$S_1$	0	1	$\frac{2}{7}$	0	$\frac{1}{7}$	16
$S_2$	0	0	$\frac{3}{7}$	1	$-\frac{2}{7}$	6
$S_3$	1	0	$-\frac{5}{7}$	0	$\frac{1}{7}$	5
$z_j - c_j$	0	0	$-\frac{8}{7}$	0	$\frac{10}{7}$	116

In this table, observe that  $x_1$  has replaced  $s_3$  in the basic column and the improved solution  $(x_1, x_2, S_1, S_2, S_3) = (5, 16, 0, 6, 0)$  has a z-value of

$$z = 4x_1 + 6x_2 = 4(5) + 6(16) = 116$$

### **Third Iteration**

In this tableau, there is still a negative entry in the bottom row. Thus, you will choose  $S_2$  as the entry variable and  $S_3$  as the departing variable, as shown in the following tableau.

**Table 3.7:**

$B$	$x_1$	$x_2$	$S_1$	$S_2$	$S_3$	$x_B$	$\theta$
$S_1$	0	1	$\frac{2}{7}$	0	$\frac{1}{7}$	16	56
$S_2$	0	0	$\frac{3}{7}$	1	$-\frac{2}{7}$	6	14
$S_3$	1	0	$-\frac{5}{7}$	0	$\frac{1}{7}$	5	-
$z_j - c_j$	0	0	$-\frac{8}{7}$	0	$\frac{10}{7}$	116	

The pivot entry is the entry in the second row and third column as shown in the table above. By performing one more iteration of the simplex method, you will obtain the following tableau.

**Table 3.8: Final Tableau**

$B$	$x_1$	$x_2$	$S_1$	$S_2$	$S_3$	$x_B$
$S_1$	0	1	0	$-\frac{2}{3}$	$\frac{1}{3}$	12
$S_2$	0	0	1	$\frac{7}{3}$	$-\frac{2}{3}$	14
$S_3$	1	0	0	$\frac{5}{3}$	$-\frac{1}{3}$	15
$z_j - c_j$	0	0	0	$\frac{8}{3}$	$\frac{10}{7}$	132

In this tableau, there are no negative elements in the bottom row. You have therefore determined the optimal solution to be  $(x_1, x_2, S_1, S_2, S_3) = (15, 12, 14, 0, 0)$  with  $z = 4x_1 + 6x_2 = 4(15) + 6(12) = 132$ .

*It is worth noting that ties may occur in choosing entering and/or departing variables. Should this happen, any choice among the tied variables may be made.*

### C. Minimization Problem

A minimization problem is in standard form if it is of the form

$$\begin{aligned}
 & \text{Minimize } w = c_1x_1 + c_2x_2 + \dots + c_nx_n \\
 & \text{Subject to } \begin{aligned}
 & a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n \geq b_1 \\
 & a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n \geq b_2 \\
 & a_{m1}x_1 + a_{m2}x_2 + \dots + a_{mn}x_n \geq b_m \dots \dots (3.6.10)
 \end{aligned}
 \end{aligned}$$

where  $x_i \geq 0$  and  $b_i \geq 0$ . The basic procedure used to solve such a problem is to convert it to a maximization problem in standard form, and then apply the simplex method as discussed above.

**Illustration** Minimization Problem.

Solve the following.

$$\begin{aligned}
 & \text{Minimize } w = 0.12x_1 + 0.15x_2 \\
 & \text{Subject to } \begin{aligned}
 & 60x_1 + 60x_2 \geq 300 \\
 & 12x_1 + 66x_2 \geq 36 \\
 & 10x_1 + 30x_2 \geq 90 \\
 & x_1, x_2 \geq 0
 \end{aligned}
 \end{aligned}$$

**Solution.** Using the simplex method, the first step in converting this problem to a maximization problem is to form the augmented matrix for this system of inequalities. To this augmented matrix you add a last row that represents the coefficients of the objective function, as follows.

$$\begin{bmatrix}
 60 & 60 & : & 300 \\
 12 & 66 & : & 36 \\
 10 & 30 & : & 90 \\
 0.12 & 0.15 & : & 0
 \end{bmatrix}$$

Next, form the transpose of this matrix by interchanging its rows and columns.

$$\begin{bmatrix}
 60 & 12 & 10 & : & 0.12 \\
 60 & 66 & 30 & : & 0.15 \\
 300 & 36 & 90 & : & 0
 \end{bmatrix}$$

Note that the rows of this matrix are the columns of the first matrix, and vice versa. Finally, interpret the new matrix as a maximization problem as follows. (To do this, we introduce new variables,  $y_1$ ,  $y_2$ , and  $y_3$ .) You call this corresponding maximization problem the dual of the original minimization problem.

### Dual Maximization Problem

$$\text{Maximize } z = 300y_1 + 36y_2 + 90y_3 \text{ Dual objective function}$$

$$\text{Subject to } 60y_1 + 12y_2 + 10y_3 \leq 0.12 \text{ Dual constraints}$$

$$60y_1 + 6y_2 + 30y_3 \leq 0.15$$

$$\text{where } y_1 \geq 0, y_2 \geq 0 \text{ and } y_3 \geq 0.$$

As it turns out, the solution of the original minimization problem can be found by applying the simplex method to the new dual problem, as follows.

**Table 3.9:**

$B$	$y_1$	$y_2$	$y_3$	$S_1$	$S_2$	$y_B$	$\theta$
$S_1$	<b>60</b>	12	10	1	0	0.12	0.002
$S_2$	60	6	30	0	1	0.15	0.004
$z_j - c_j$	300	36	90	0	0	0	

$B$	$y_1$	$y_2$	$y_3$	$S_1$	$S_2$	$y_B$	$\theta$
$S_1$	1	$\frac{1}{5}$	$\frac{1}{6}$	$\frac{1}{60}$	0	$\frac{1}{500}$	$\frac{3}{250}$
$S_2$	0	-6	<b>20</b>	-1	1	$\frac{3}{100}$	$\frac{3}{2000}$
$z_j - c_j$	0	24	40	5	0	$\frac{3}{5}$	

$B$	$y_1$	$y_2$	$y_3$	$S_1$	$S_2$	$y_B$	$\theta$
$S_1$	1	$\frac{1}{4}$	0	$\frac{1}{40}$	$\frac{1}{120}$	$\frac{7}{4000}$	$\frac{3}{250}$
$S_2$	0	$-\frac{3}{10}$	1	$\frac{1}{20}$	$\frac{1}{20}$	$\frac{3}{2000}$	$\frac{3}{2000}$
$z_j - c_j$	0	12	0	$3 = x_1$	$2 = x_2$	$\frac{33}{50}$	

As it turns out, the solution of the original minimization problem can be found by applying the simplex method to the new dual problem, as follows.

Thus, the solution of the dual maximization problem is  $z = \frac{33}{50} = 0.66$ . This is the same value you obtained using graphical method. The  $x$ -values corresponding to this optimal solution are obtained from the entries in the bottom row corresponding to slack variable columns. In other words, the optimal solution occurs when  $x_1 = 3$  and  $x_2 = 2$ .

The fact that a dual maximization problem has the same solution as its original minimization problem is stated formally in a result called the von Neumann Duality Principle, after the American mathematician John von Neumann (1903-1957).

**The von Neumann Duality Principle:** The objective value  $w$  of a minimization problem in standard form has a minimum value if and only if the objective value  $z$  of the dual maximization problem has a maximum value. Moreover, the minimum value of  $w$  is equal to the maximum value of  $z$ .

### SELF ASSESSMENT EXERCISE

Solve graphically the following LPP.

$$\begin{aligned} \text{Maximize } z &= 20x_1 + 10x_2 \\ \text{Subject to } x_1 + 2x_2 &\leq 40 \\ 3x_1 + x_2 &\geq 30 \\ 4x_1 + 3x_2 &\geq 60 \\ \text{with } x_1, x_2 &\geq 0 \end{aligned}$$

### 4.0 CONCLUSION

In this unit, we studied the graphical and the algebraic methods for solving a linear programming problem. We have also seen their limitations. With these limitations of the algebraic method, it becomes imperative to consider a method that is better than the algebraic method and the graphical method. This method

- i. Would not evaluate infeasible solutions.
- ii. Should progressively give you better solutions.
- iii. Should be able to terminate as soon as it has found the optimum. It should not put you in a situation where you have evaluated the optimum but still have to evaluate the rest before you would realize that you have arrived at an optimum solution earlier.

Also, a method that can do all these would add more value to the algebraic method that you have seen. Obviously, that method would require more computation and extra effort. This method is called the simplex method which is essentially an extension of the algebraic method and exactly addresses the three concerns you have listed above. Simplex method is the most important tool that had been developed to solve linear programming problems.

### 5.0 SUMMARY

Having gone through this unit, you are now able to;

- (i) Solve linear programming problems using graphical methods
- (ii) Solve linear programming problems using algebraic methods.

- (iii) Any feasible solution to LPP, which satisfies the non-negativity restriction is called its feasible solution.
- (iv) Any feasible solution, which optimizes (minimizes or maximizes) the objective function of the LPP is called optimum solution.
- (v) If the value of the objective function can be increased or decreased indefinitely, such solutions are called unbounded solutions.
- (vi) Know how to solve linear programming problem using the algebraic and tabular simplex algorithms.

## 6.0 TUTOR MARKED ASSIGNMENT

Use the simplex method to solve the given linear programming problem.

1. Maximize  $z = 2x_1 + 3x_2 + 5x_3$   
 Subject to:  $3x_1 + 10x_2 + 5x_3 \leq 15$   
 $33x_1 - 10x_2 + 9x_3 \leq 33$   
 $x_1 + 2x_2 + x_3 \geq 4$   
 with  $x_1, x_2, x_3 \geq 0$
  
2. Minimize  $w = 3x_1 + 3x_2$   
 Subject to:  $2x_1 + x_2 \geq 15$   
 $x_1 + x_2 \geq 12$   
 with  $x_1, x_2 \geq 0$
  
3. Minimize  $w = 2x_1 + 2x_2$   
 Subject to:  $x_1 + 2x_2 \geq 3$   
 $3x_1 + 2x_2 \geq 5$   
 with  $x_1, x_2 \geq 0$

## 7.0 REFERENCES/FURTHER READING

- Denardo, E. V. (2002). *The Science of Decision Making: A Problem-Based Approach Using Excel*. New York: John Wiley
- Dixon-Ogbechi, B.N. (2001). *Decision Theory in Business*. Lagos: Philglad Nig. Ltd.
- Dwivedi, D. N. (2002). “*Linear Programming*” *Managerial Economics*. (6th ed.). New Delhi: Vikas Publishing House PVT Ltd.
- Francis, A. (1998). “*Linear Equations*” *Business Mathematics and Statistics*. (5th ed.). England: Ashford Colour Press.
- Gupta, P.K. & Hira, D.S. (2012). *Operations Research*. New-Delhi: S. Chand & Company.
- Linear Programming Lectures Youtube videos
- Lucey, T. (1988). *Quantitative Techniques: An Instructional Manual*. London: DP Publications.
- Luenberger D. G (1973). *Introduction to Linear and Nonlinear Programming*, Addison-Wesley, Reading, Mass.
- Mokhtar S. B., John, J. J. (1977). *Linear Programming and Network Flows*, John Wiley and Sons, New York.
- Nicholson, W. (1978). *Microeconomic Theory: Basic Principles and Extensions*. (2n ed.). Illinois: The Dryden Press.
- Shetty, C. M. (1959). Solving Linear Programming Problems with Variable parameters, *Journal of Industrial Engineering*, Vol. 10(6).